

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: October 12, 2004, 15:06:21 ; Search time 30 Seconds
(without alignments)
7158.527 Million cell updates/sec

Title: US-10-046-935-2234
Perfect score: 1962
Sequence: 1 atggacgtcgccggtgccc.....aatttgaatgaagcataa 1116

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlp
-O=/cgn2_1/USPTO_spool_p/US10046935/runat_12102004_155226_15255/app query.fasta_1.1287
-DB=pir_79 -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10046935 @CGN_1_1_46 @runat_12102004_155226_15255 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

pir_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	602.5	30.7	202	T50635	hypothetical prote
2	289.5	14.8	462	D85438	hypothetical prote
3	250.5	12.8	515	F09700	protein F12A21.9 [
4	170	8.7	555	T01142	hypothetical prote
5	146	7.4	851	D86254	hypothetical prote
6	142	7.2	950	F86222	hypothetical prote
7	132.5	6.8	570	T49181	cyclophilin-like p
8	131.5	6.7	906	T01440	hypothetical prote
9	129	6.6	754	JCS314	CDG28/cdc2-like ki
10	117.5	6.0	595	T51747	RNA helicase RH26
11	115.5	5.9	1701	T43213	ENBP1 protein - ba
12	115	5.9	1051	G59436	KIAA1304 protein [
13	114	5.8	691	T48506	hypothetical prote
14	113	5.8	1641	T10955	early nodulin bind

15	111.5	5.7	1711	1	A47392	chromodomain-helic
16	107.5	5.5	395	2	T19144	hypothetical prote
17	107	5.5	561	2	T14792	hypothetical prote
18	106	5.4	447	2	T20552	hypothetical prote
19	105.5	5.4	857	2	T05352	hypothetical prote
20	105.5	5.4	1634	2	T26517	hypothetical prote
21	105	5.4	1200	2	T00042	BH-protocadherin p
22	105	5.4	1746	1	S19694	tenascin precursor
23	104.5	5.3	770	2	H84463	hypothetical prote
24	104.5	5.3	855	2	S56948	HAU5 protein - yea
25	104	5.3	519	2	S23796	modifier-3 protein
26	103.5	5.3	891	2	G84693	probable proline-r
27	103.5	5.3	1203	2	S26650	DNA-binding protei
28	103.5	5.3	1216	2	T09224	spindle assembly c
29	102.5	5.2	706	2	S43581	C28A5.6 protein (c
30	102.5	5.2	1061	2	G88428	protein C28A5.6 [i
31	102	5.2	560	2	T29586	hypothetical prote
32	102	5.2	925	2	T16235	hypothetical prote
33	101.5	5.2	451	2	T16418	hypothetical prote
34	101.5	5.2	932	2	T45894	hypothetical prote
35	101.5	5.2	1016	2	T25433	hypothetical prote
36	101.5	5.2	1269	2	T00443	hypothetical prote
37	101.5	5.2	1389	2	H84914	probable WD-40 rep
38	101	5.1	454	2	E84524	En/Spm-like transp
39	101	5.1	1252	2	T00263	hypothetical prote
40	101	5.1	1791	2	T02345	hypothetical prote
41	100.5	5.1	362	2	B64527	M-protein - Helico
42	100.5	5.1	553	2	S22954	finger protein zip
43	100.5	5.1	887	2	T39492	conserved hypothet
44	100.5	5.1	1289	2	T00387	KIAA0622 protein -
45	100.5	5.1	1307	2	G96711	unknown protein, 9

ALIGNMENTS

RESULT 1

T50635
hypothetical protein DKFZp762L0311.1 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 09-Jul-2004

C:Accession: T50635

R:Blum, H.; Bauersachs, S.; Newes, H.W.; Weil, B.; Wiemann, S.

submitted to the Protein Sequence Database, June 2000

A:Reference number: Z25145

A:Accession: T50635

A>Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-202 <AAA>

A:Cross-references: UNIPROT:Q9NPV2; EMBL:AL359619

A:Experimental source: adult melanoma (MeWo cell line); clone DKFZp762L0311

C:Genetics:

A:Note: DKFZp762L0311.1

Alignment Scores:

Pred. No.:	8.68e-44	Length:	202
Score:	602.50	Matches:	116
Percent Similarity:	69.23%	Conservative:	28
Best Local Similarity:	55.77%	Mismatches:	47
Query Match:	30.71%	Indels:	17
DB:	2	Gaps:	4

US-10-046-935-2234 (1-1116) x T50635 (1-202)

QY	481	AGGCGTACATCCCG-----GGTGTCTCCAGGAGAAACCTCGAACGGAGACTCGT	534
Db	5	ArgArgAlaPheSerGluGlyGlnIleThrArgArgMetAsnProThrArgSerAlaArg	24
QY	535	CCTCTTACCAGGTCAAGGTCCCGGATCTCGGGTCCCTTGACGCTCTACCATGGRGAGA	594
Db	25	ProProGluLysPhe-----AlaLeuGluAsnPheThrValSerAla	38
QY	595	GCCAGTTTKGAKMTMWTGATCATGTGTGTGAGAAAGAGACCGTGGATGGCTACATG	654

[illegible]

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Nature 408, 816-820, 2000

A;Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
A;Reference number: A84420; MUID:20083487; PMID:106117197

A;Accession: H84625
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-555 <STO>
A;Cross-references: GB:AE002093; NID:g3242716; PIDN:AAC23768.1; GSPDB:GN00139
C;Genetics:
A;Gene: F26B6.18; At2g23530
A;Map position: 2
A;Introns: 87/2; 166/2; 188/1; 231/2
C;Superfamily: Arabidopsis thaliana hypothetical protein F26B6.18

Alignment Scores:
Pred. No.: 2,34e-13 Length: 515
Score: 250.50 Matches: 48
Percent Similarity: 58.49% Conservative: 14
Best Local Similarity: 45.28% Mismatches: 37
Query Match: 12.77% Indels: 7
DB: 2 Gaps: 2

US-10-046-935-2234 (1-1116) x F96700 (1-515)

QY 778 AGATATATAACCGTTCACGTGGCTCTACTGTGATCAATCGCGTCAGAGACTATTGAT 837
:::|||||:::|||||
Db 25 ArgTyrAspSerSerAsnGlyLysThrCysHisGlnCysArgGlnLysThrMetAsp 44
:::|||||:::|||||

QY 838 ACCAAACAAACATGC-----AGNAACCCAGACTGCTGGGGCGTTCAGGCCAG 895
:::|||||:::|||||
Db 45 PheValAlaSerCysLysAlaMetLysLysAspLysGlnCys-----ThrIleAsn 61
:::|||||:::|||||

QY 886 TTCTGTGGCCCTCGCTTCGAAACCGTTATGTGTGAAGAGGTGAGGATGCTCTGCTGGAT 945
|||||
Db 62 PheCysHisLysCysLeuIleAsnArgTyrGlyLysAlaGluValAlaLysLeu 81
:::|||||:::|||||

QY 946 CGGAATCGCATGTCGGCGCTTCGAGGAATCTGCAACTGCAATGCTGCGCGGACGCA 1005
:::|||||:::|||||

Db 82 AspAspTyrLysCysProGlnCysArgGlyIleCysAsnCysSerPheCysArgLysLys 101
:::|||||:::|||||

QY 1006 GATGGACGGTGTGCGACTGGGGCTGTGTGATTAGCCAAATATCATGCTTTGGGAAT 1065
|||||
Db 102 ArgGlyLeuAsnProThrGlyIleLeuAlaHisLysAlaLysAlaSerGlyLeuAlaSer 121
:::|||||:::|||||

QY 1066 GTGCATGCTCTACTTGAAA 1083
|||||
Db 122 ValSerMetLeuLeuGlu 127
:::|||||:::|||||

RESULT 4
T01142
hypothetical protein At2g23530 [imported] - Arabidopsis thaliana
N;Alternate names: hypothetical protein F26B6.18
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C;Accession: T01142; H84625
R;Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, submitted to the EMBL Data Library, June 1998
A;Description: Arabidopsis thaliana chromosome II BAC F26B6 genomic sequence.
A;Reference number: Z14198
A;Accession: T01142
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-555 <ROU>
A;Cross-references: UNIPROT:O80469; EMBL:AC003040; NID:g3242700; PID:g3242716
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VarAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
A;Reference number: A84420; MUID:20083487; PMID:106117197

A;Accession: H84625
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-555 <STO>
A;Cross-references: GB:AE002093; NID:g3242716; PIDN:AAC23768.1; GSPDB:GN00139
C;Genetics:
A;Gene: F26B6.18; At2g23530
A;Map position: 2
A;Introns: 87/2; 166/2; 188/1; 231/2
C;Superfamily: Arabidopsis thaliana hypothetical protein F26B6.18

Alignment Scores:
Pred. No.: 2,1e-06 Length: 555
Score: 170.00 Matches: 78
Percent Similarity: 40.81% Conservative: 33
Best Local Similarity: 28.68% Mismatches: 92
Query Match: 8.66% Indels: 69
DB: 2 Gaps: 12

US-10-046-935-2234 (1-1116) x T01142 (1-555)

QY 274 CGCCAGCCCTCAGAGAAATCTGTGACTGATTCCAACTCCGATTTCAGAA-----GAT 324
|||||
Db 2 ArgThrGluAlaGlnAspSerValProLysSerAsnProAsnProGluLeuIleLysGlu 21
:::|||||:::|||||

QY 325 GAAAGTGAATGAATTTTGGAGAAA---AGGGCTTTAAATATAAAGCAAAACAAAGCA 381
:::|||||:::|||||

Db 22 ThrProLysValSerLeuTyrGluGlnCysArgGluGluArgLysGluAsnLeuGln 41
:::|||||:::|||||

QY 382 ATGCTTGCAAACTCATCTCTGATTTAGAAAGCTTCCCTGGCTCTCCGTGGAAGACAT 441
:::|||||:::|||||

Db 42 ArgMetAsnAsnLeu-----GlyLeuLeu 49
:::|||||:::|||||

QY 442 CCCTCCCGAGGTCCGACTCACAATCAAGAGACCCGAGGCGTACATTCCCGGGTGT 501
|||||
Db 50 AsnLeuSerArgLysLeuLysProLysThrArgProValLysArgSerTyrGlyAsnArg 69
:::|||||:::|||||

QY 502 GCTTCCAGAGAAACCTCGAAGAGAGCTGCTCTTACCAGGTCAAGGTCCCGGATC 561
:::|||||:::|||||

Db 70 AsnSerValGlnAsnProThrProLeuGlnProSerArgArgSer---SerArgLeu 88
:::|||||:::|||||

QY 562 CTCGGTCTCTGACCTCTACCATCGAGAGCCAGTTCGARKMTWGTACATGTTG 621
:::|||||:::|||||

Db 89 GluAsnThrThrProValIle----- 95
:::|||||:::|||||

QY 622 GTGAGAAAGAGGAGACCCGTGGATGGCTACATGAATGAATGATGACCTGCCAGAGCCGT 681
|||||
Db 96 -----TyrThrAspGly---IleAsnGluLysGlyLysAlaSerLys 109
:::|||||:::|||||

QY 682 CGCTCCAGATCATCGTGACCTTCGCGATATAATTGCGCCAGTGGAGAAATTACAGAG 741
|||||
Db 110 -----ArgLysSerValValIleGlyGluGlyIleArg-----AlaGluIleTyrThr 125
:::|||||:::|||||

QY 742 GAGGAGTTCGAGAACCTGCGACCAATTCGNA----- 774
|||||
Db 126 GluGluHisLysLeuLeuGlyAsnThrGluArgSerTyrThrCysPheValAspGly 145
:::|||||:::|||||

QY 775 -----GAGAAGATATATAACCGTTCACTGGGCTCTACTTGTTCATCAATGC 819
:::|||||:::|||||

Db 146 TyrAspLysAsnGlyLysArgIleTyrAspProPheAsnGlyLysThrCysHisGlnCys 165
:::|||||:::|||||

QY 820 CGTCAGAACACTATTGATACCAAAACAAACTGCAGAAACCCAGACTGCTGGGCGGTCGA 879
|||||

Db 166 ArgGlnLysThrMetGlyHisArgThrGlnCys-----SerGluCysAsnLeuValGln 183
:::|||||:::|||||

QY 880 GCCCAGTTCTGT-----GGCCCCCTCCCTCGNAACCGTTATGTCGAGAGATCTGCAACTGCACT 930
|||||

Db 184 GlyGlnPheCysAspMetValAsnMetCysLeuLys----- 195
:::|||||:::|||||

QY 931 GATGCTCTCTGGATCCGAACTGGCATTCGCCCGCTTGTGCGAGGAATCTGCAACTGCACT 990
:::|||||:::|||||

```
Db 196 ---LeuLeuArg-ileGlnThrGlyPheValProLeuValAlaGlySerAlaIleAlaVa 214
QY 991 TTCTCCGGCAGGAGATGCGCGGTGGGACTG 1024
Db 214 lCysAlaGlyIleIleLysAspGlyPheGlnLeu 225
RESULT 5
D86254
hypotheical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: D86254
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: AB6141; MUID:21016719; PMID:11130712
A:Accession: D86254
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-851 <STO>
A:Cross-references: UNIPROT:065394; GB:AE005172; NID:g3157933; PIDN:AAC17616.1; GSPDB:GN
A:Genetics:
A:Map position: 1
Alignment Scores:
Pred. No.: 0.00261 Length: 851
Score: 146.00 Matches: 72
Percent Similarity: 40.40% Conservative: 50
Best Local Similarity: 23.84% Mismatches: 128
Query Match: 7.44% Indels: 52
DB: 2 Gaps: 17
US-10-046-935-2234 (1-1116) x D86254 (1-851)
QY 268 GAGTCCCGCAGCCCTCAGAG---AATCTGTGATGATTCGAACTCCGATTCGAGAT 324
Db 28 GluLysLysLysProValGluAlaThrSerLeuSerSerGlySerSerAspIleGluGlu 47
QY 325 GAAAGTGGATGAATTTTGGAGAAAGCGCTTAAATATAAGCAACAAACAAAGCAATG 384
Db 48 GluIleSerValGlu---CysProLysArgValAlaAsnGlnArgLysArgSerLys 66
QY 385 CTTGCAAACTCATGCTCTGAATAGAAAGCTTCCCTGGCTCGTTCGTCGGAAGACATCC 444
Db 67 AlaAspGluIleLysThrLys-----SerSerArgLysArgLysCys 80
QY 445 CTCACGAGCTCGCATCAACAATCAAGAGACCGGAGCGGTACATCCCGGGTGTGCT 504
Db 81 AspAspGluAsnLysCysGluGluAsnGluLysLysGlnArgSer-----SerValLys 98
QY 505 TCCAGAGAAACCTCAACGAGAGCTCGTCTCTTACCAGGTCAGGTCGCGATCCTC 564
Db 99 LysArgAlaThrThrTrpLysGluGluGluValValAspAspGluLysLysCysGlu 118
QY 565 GGGTCCCTTACGCTTACCTACCGAGAGAGCGAGTTCGARKMTWMTGATCTGTGGTG 624
Db 119 GlnGlnLeuGlnLeuValProSerSerLysAlaThrSerArgSerArg-----Ser 135
QY 625 AGAAGAGAAACCGTGTGATGCTACATG-----AATGAGATGACCTGCCAGA---- 675
Db 136 LysLysSerValSerValAspThrTrpLeuValAsnAsnGluIleAspValSerAlaLeu 155
QY 676 AGCCGTCGCTCCAGATCATCCGTGACCTTCGCGATATAATTCGCCCGAGGAATTT 735
Db 156 SerSerArgSerGluSerGluLeuSer-----AspSerTyr 167
```

```
QY 736 ACAGAGGAGGAGTTGTCAGAACGTCCTGAGCAATTTCTCGAGAGACAGATATATAACCGTTCA 795
Db 168 LeuLysThrGluTyrPheAsnAspCys---ArgSerMetThrArgSerLeuLysAlaAsn 186
QY 796 CTGGGC-----TCTACTTGTCAATCAATGC-----CGTCAGAGACTATTGTATACC 840
Db 187 LeuGlyGluLeuAlaIleCysHisGlnCysSerLysGlyGluArgArgTyrLeuPheIle 206
QY 841 AAACAAACTGCAGAAACCCAGACTGCTGGGGCGTTCGAGGCCACTTCTGTGGCCCTCGC 900
Db 207 CysThrPheCysGlu-----ValArg---LeuTyrCysPheProCys 219
QY 901 CTTCAAAACCGTATGTGTGAAGAGTCAGGGATGCTCTGCTGGATCCGAACATCGCATTC 960
Db 220 IleLysLysTyrTyrProHisLeuSerThrAspAspIleLeuGlu-----LysCys 236
QY 961 CGCCCTTTCGAGGAATTCGAACTGTCAGTTCCTCCGCGCAGCAGATGAGCGGTGTGG 1020
Db 237 ProPheCysArgGlyThrCysAsnCysCysThrCysLeuHisSerSerGlyLeuIleGlu 256
QY 1021 ACTGGGGTCTGTGTTATTATAGCCAAATATCATGCTTTGGGAATGTGCAT----- 1071
Db 257 ThrSerLysArg---LysLeuAspLysTyrGluArgPheTyrHisLeuArgPheLeuIle 275
QY 1072 -----GCCTACTTCAAAAGCCTG-----AAACAGGAATTTGAAATG 1107
Db 276 ValAlaMetLeuProPheLeuLysLysLeuLysLysAlaGlnAspGlnGluIleGluThr 295
QY 1108 CAAGCA 1113
Db 296 GluAla 297
RESULT 6
F86222
Hypotheical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: F86222
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: AB6141; MUID:21016719; PMID:11130712
A:Accession: F86222
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-950 <STO>
A:Cross-references: UNIPROT:004024; GB:AE005172; NID:g2342679; PIDN:AAB70402.1; GSPDB:GN
A:Genetics:
A:Map position: 1
Alignment Scores:
Pred. No.: 0.000585 Length: 950
Score: 142.00 Matches: 78
Percent Similarity: 32.15% Conservative: 31
Best Local Similarity: 23.01% Mismatches: 128
Query Match: 7.24% Indels: 102
DB: 2 Gaps: 15
US-10-046-935-2234 (1-1116) x F86222 (1-950)
QY 103 GATGACAGTGTGACAGCTTTCCTTCTGTGATAATTTTGGCAACACGAGGTCAGTCAGTT 162
Db 6 AspGluThrCysAspSerVal----- 12
QY 163 CGGAGAGGCTGTAGGACCGCAGCCAGTCAGGCATCTGGACCTCTCAGGGTGGCGATG 222
```

Db 13 -----leuArgMetAsnAla 17
 QY 223 AAGTTTCCAGCGGAGTACCAGGGAGAGCAACCAACAAAAGCAGAGTCCCGCCAGCCC 282
 Db 18 AsnGluGlnThrArgSerAlaAsnGlyIleGlyAsnGlyAsnGlySerIleProGly 37
 QY 283 TCAGAGAATTCTGTGACTGATTCCCAACTCCGATTTCAGAGAT----- 324
 Db 38 IleProAspLeuArgCysLysArgSerAspGlyLysGlnTrpArgCysThrAlaMet 57
 QY 325 -----GAAAGTGGGAATGAATTTTGGAGAAAAGGGCTTTA 360
 Db 58 SerMetAlaAspLysThrValCysGluLysHisTyrIleGlnAlaLysLysArgAlaAla 77
 QY 361 AAT-----ATAAGCAAAACAAAGCAATCTTGCAAACTCATGCTGAATTA 408
 Db 78 AsnSerAlaPheArgAlaAsnGlnLysLysAlaLysArgSerSerLeuGlyGluThr 97
 QY 409 GAAAGCTTCCCTGGCTCGTTCCGTGGAAGACATCCCTCCCA----- 450
 Db 98 AspThrTyrSerGluGlyLysMetAspAspPheGluLeuProValThrSerIleAspHis 117
 QY 451 -----GCTCCGACTCACAATCAAGAGAGACCGCGAAGCGGTACA 489
 Db 118 TyrAsnAsnGlyLeuAlaSerAlaSerLysSerAsnGlyArgLeuGluLysArgHisAsn 137
 QY 490 TTCCCGGGTGTCTTCCAGGAGAACCTCGACGAGAGCTCGTCTCTTACCAGTCA 549
 Db 138 -----LysSerLeuMetArgTyrSerProGlu-----ThrProMetMetArgSer 152
 QY 550 AGGTCC---CGGATCCTCCGGTCCCTT---GACGCTCTACCATCTGGRGAGAGCCAGTTTK 603
 Db 153 PheSerProArgValAlaValAspLeuAsnAspAspLeuGlyArgAspValMetPhe 172
 QY 604 GARKMTWGTAGATGTGTGTGAGAAAGAGAGACCGTGGATGGCTACATGAATGAAGAT 663
 Db 173 GluGluGlyTyrArgSerTyrArgThrProProSerVal---AlaValMetAsp----- 189
 QY 664 GACCTGCCAGAGCGCTCGCTCCAGATCATCGTGACCTCCGATATATTCGCCCA 723
 Db 190 -----ProThrArgAsnArgSerHisGlnSerThrSer-----Pro 201
 QY 724 GTGGAAGAAATTACAGAGGAGGAGTTGGAGAACGCTCTGCAGCAATCTCGAGAGAAGATA 783
 Db 202 MetGlu-----TyrSerAlaAlaSerThrAspVal 211
 QY 784 TATAACCGTTCACTGGCTCTACTTGTCTATCATGTCGTCAGAGACTATTGATACCAA 843
 Db 212 SerAlaGluSerLeuGlyGluIleCysHisGlnCysGlnArgLysAspArgGluArgIle 231
 QY 844 ACAAACTGCAGAAACCCAGCTCTCGTGGCGGTTCGAGGCCAGTCTGTGGCCCTCCCTT 903
 Db 232 IleSerCysLeuLysCysAsn-----GlnArgAlaPheCysHisAsnCysLeu 247
 QY 904 GAAACCGTTATGGT-----GAAGAGTTCAGGGAGTCTGTGTCGATCCGAAC 951
 Db 248 SerAlaArgTyrSerGluIleSerLeuGluValGluLysVal----- 262
 QY 952 TGGCATTCGCGCTTGTGAGGAATCTGCAACTCGAGTTTCTGCGGAGGAGAT 1008
 Db 263 -----CysProAlaCysArgGlyLeuCysAspCysLysSerCysLeuArgSerAsp 279

RESULT 7

T49181
 cyclophilin-like protein - Arabidopsis thaliana
 N;Alternate names: protein MAA21.30
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
 C;Accession: T49181
 R;Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Rudd, S.; Lemcke, K.
 submitted to the Protein Sequence Database, April 2000
 A;Reference number: Z25018

A;Accession: T49181
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-570 <RIB>
 A;Cross-references: UNIPROT:O9LY75; EMBL:AL163818; GSPDB:GN00061; ATSP:MAA21.30
 A;Experimental source: cultivar Columbia; BAC clone MAA21
 C;Genetics:
 A;Gene: ATSP:MAA21.30
 A;Map position: 3
 A;Introns: 26/3; 45/1; 74/3; 83/1; 96/1; 135/2; 189/1; 341/2; 455/2; 482/2
 F;6-175/Domain: cyclophilin homology <CYP>

Alignment Scores:
 Pred. No.: 0.00363 Length: 570
 Score: 132.50 Matches: 79
 Percent Similarity: 39.09% Conservative: 41
 Best Local Similarity: 25.73% Mismatches: 108
 Query Match: 6.75% Indels: 80
 DB: 2 Gaps: 15

US-10-046-935-2234 (1-1116) x T49181 (1-570)

QY 4 GACGCTCGCGGTCCGCGAGAAAGATCTCAGAGTAAAGAAAGAACTTAAAGAAATTCAGA 63
 Db 207 GluAlaLysGluThrArgLysLysGluSerAsnGluLysArgLysArgLysArg 226
 QY 64 TATGTCAAGTTGATTTCATGGAACCTCGTCATCTCTGATGACAGATTGTGACAGC--- 120
 Db 227 TyrSerSerSerAspSerTyrSerSerSerSerSerSerSerSerSerSerSer 246
 QY 121 ---TTGCTTTCGATAATTTTGCAAACACGAGCTGCAGTTCGCGAAGGCTGTAG 177
 Db 247 AlaTyrSerSerSerSerTyrGluSerSerSerSerSerSerSerSerSerSer 266
 QY 178 ACCGCGACGAGTCGAGGCATCTCTGACCTCTCAGGGTGGCGATGAAGTTTCAGCGCG 237
 Db 267 ---LysSerThrThrArgHisLysGly-----ArgArgGlyGluArg 279
 QY 238 AGTACCAGGGGACCAACCAACAAA---GCAGAGTCCGCGACCCCTCAGAGAT 291
 Db 280 LysSerLysGlyArgSerGlyLysLysLysAlaArgProAspArgLysProSerThr 299
 QY 292 TCTGTG-----ACTGATTCCCAATCCGATTCCAGAGATGAAAGTGAATGAATTTT 345
 Db 300 SerSerSerAspThrGluSerSerSerSerSerSerSerSerSerSerSer 315
 QY 346 GAGAAAGGGCTTTAAATATAAAGCAAAACAAAGCAATCTCTGCAAAATCATGTCTGAA 405
 Db 316 GlyHisLysAla-----IleLysSerValLysValAspAsnAlaAspGlnHisAla 333
 QY 406 TTAGAAAGCTTCCCTGGCTCGTTCGTTGGAAGACATCCCTCCAGGCTCCGACTCACAA 465
 Db 334 LeuAspAsp-----SerValLysSerArg 341
 QY 466 TCAAGAGACCCGAGGCGTACATTCGCGGTGTGTCTTCAGAGAGAAACCTGAACCG 525
 Db 342 SerArgSerProIleArg-----ArgArgAsnGlnAsnSer 353
 QY 526 AGAGCTGTCTCTTACACAGGTCAAGTCCCGATCTCCGGTCCCTTGCAGCTCTACCC 585
 Db 354 ArgSerLysSerProSerArgSerProValArgValLeuGlyAsnGlyAsnArgSerPro 373
 QY 586 ATGGRGAGCCAGTTTGGATGATGTTGGTGCAGAAAGAGGAGCCCTGGAT 645
 Db 374 -----SerArgSerProValArg 379
 QY 646 GGCTACATGAATGA-----AGATGACTGCCAGAGCCGCTCCATCATCGTGAC 701
 Db 380 AspLeuGlyAsnGlySerArgSerProArgGluLysProThrGluGluThrValGlyLys 399
 QY 702 CTTCCGCATATAATTCGCCAGTGA-----AGAAATTACAGAGAGGAGTGGAGAA 755
 Db 400 SerPheArgSerProSerProSerGlyValProLysArgIleArgLysGlyArgGly--- 418

QY 756 CGTCTGCAGCAATTCCTCAGAGAGAGATATATAACCGTTTCACTGCGCTCTACTTGTGTCATCA 815
Db 419 -----PheThr----- 420
QY 816 ATGCGCTCAGAGACTATTGATACCAAAACAACTGCAGAAACCCAGACTGCTGGGGCGT 875
Db 421 -----GluArgTyrSerPheAlaArgLysTyrHisThrProSer---ProGluArg 436
QY 876 TCGAGGCCAGTCTCTGTGGCCC 896
Db 437 SerProProArgHisTrpPro 443
RESULT 8
T01440
hypothetical protein F2401.2 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C:Accession: T01440
R:Shinn, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.; Sun, H.; Conway, A.; Con-
eologis, A.; Ecker, J.R.
submitted to the EMBL Data Library, January 1998
A:Description: Genomic sequence for Arabidopsis thaliana BAC F2401.
A:Reference number: Z14211
A:Accession: T01440
A>Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-906 <SHI>
A:Cross-references: UNIPROT:048794; EMBL:AC003113; NID:g2689438; PID:g2781345; GSPDB:GNO
A:Gene: ATSP:F2401.2
A:Map position: 1
A:Introns: 239/2; 272/3; 317/1; 338/2; 682/3; 832/1; 860/3
Alignment Scores:
Pred. No.: 0.00469 Length: 906
Score: 131.50 Matches: 70
Percent Similarity: 33.87% Conservative: 36
Best Local Similarity: 22.36% Mismatches: 90
Query Match: 6.70% Indels: 117
Gaps: 14
US-10-046-935-2234 (1-1116) x T01440 (1-906)
QY 253 ACCAACAACAAAGCAGAGTCCCGCAGCCC-----TCAGAGAATTCTGTGACTGATCC 306
Db 22 ThrThrArgLysLysArgThrLeuLysProPheMetSerLysGlySerSerProSer 41
QY 307 AACCTCCGATTCA-----GAAGATGAAGTGAATG 336
Db 42 SerSerAspSerArgLysArgLysLeuSerArgAlaGluAspSerAspSerAlaVal 61
QY 337 AATTTTGGAGAAAGGCTTAATATA----- 366
Db 62 LysArgAsnAlaLysArgArgLysLysLysValGluGluTyrTyrGluAspAsp 81
QY 367 -----AAGCAAAAC 375
Db 82 AspCysIleLeuSerAspTrpValGlnArgAsnThrAlaLysArgIleAspLysArgAsn 101
QY 376 AAGCAATGCTTGCAAACTCATGTCTGAATAGAAAGCTTCCCTGGCTGCTCCGTGA 435
Db 102 GluGluVal-----GluValMetValLysIleGluSer----- 112
QY 436 AGACATCCCTCCAGGCTCCGACTCACAATCAAGGAGACCGCAAGGGGTACATTCCCG 495
Db 113 -----GlyAspAspCysThrIle-----GlyLysTrpPheSer 123
QY 496 GGTGTGCTTCAGGAGAAACCTCGAACGAGAGCTCGT----- 534
Db 124 AspValSerSerLysArgLysAspLysArgGlnValGluValAspGluAspGluTrp 143
QY 535 -----CCTCTTACCAGTCAAGGTCCTCCGATCCTC 564

Db 144 GluGluGluValThrLeuCysSerLysIleLysAlaThrSerSerArgSerArg---Thr 162
QY 555 GGGTCCCTTTCAGCTCTACCCATGGAGAGACCCAGTTTGARKMTWGTACATGTTGGTG 624
Db 163 HisSerLeuSerAla----- 167
QY 625 AGAAAGAGGAAGACCGTGGTGGCTTACATGAATGAAGATGACCTGCCAGAAAGCCGCTGC 684
Db 168 -----AsnSerProGluAsnValThrAspValIleSerProCysArg 181
QY 685 TCCAGATCATCCGTGACCCCTTCCGCATATAATTCGCCAGTGGAGAAATTCAGAGGAG 744
Db 182 SerArgSer-----ProAlaSerAsnValSerAspSer 192
QY 745 GAGTTGGAGAACGCTTCGACGAATTCGAGAGAGATATATAACCGTTCCTACTGGGCTCT 804
Db 193 IleGlnLysAsnAspCysThrSerArg-----LysGlnSerGlyPro 207
QY 805 ACTTGTCATCAATCCGCTCAGAAAGACTATTGATACCAAAACAAACTGCAGAAACCCAGAC 864
Db 208 IleCysHisGlnCysLeuLysGlyGluArgIleThrLeuLeuLeuCys-----SerGlu 225
QY 865 TGCTGGGGCGTTCGAGGCCAGTTCGTGGCCCTGCTCGAAACCCGTTATGTTGTAAGAG 924
Db 226 Cys-----GluLysThrMetPheCysLeuGlnCysIleArgLysTrpTyrProAsnLeu 243
QY 925 GTCAGGATGCTCTGCTGATCCGACCTGCGATTCGCCCTGTCGAGGAATCTGCAAC 984
Db 244 SerGluAspAspValValGlu-----LysCysProLeuCysArgGlnAsnCysAsn 260
QY 985 TGCAGTTTCTCGCGCAGCAGATGACGCTGTGCGACT 1023
Db 261 CysSerLysCysLeuHisLeuAsnGlyLeuIleGluThr 273
RESULT 9
JC5314
CDC28/cdc2-like kinase associating arginine-serine cyclophilin - human
N:Alternate names: CARS-Cyp
C:Species: Homo sapiens (man)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: JC5314; G02262
R:Neatel, F.P.; Colwill, K.; Harper, S.; Pawson, T.; Anderson, S.K.
Gene 180, 151-155, 1996
A:Title: RS cyclophilins: Identification of an NK-TR1-related cyclophilin.
A:Reference number: JC5314; MUID:97128820; PMID:8973360
A:Accession: JC5314
A:Molecule type: mRNA
A:Residues: 1-754 <NES>
A:Cross-references: UNIPROT:Q13427; EMBL:U40763; NID:g1117967; PIDN:AAB40347.1; PID:g111
A:Experimental source: thymus
A:Note: submitted to the EMBL Data Library, November 1995
C:Comment: This protein contains a cyclophilin-related domain, two 140K nuclear phosphop
splicing by binding to splicing factors containing serine-arginine repeats protein.
C:Genetics:
A:Gene: GDB:CYP; CARS-CYP
A:Cross-references: GDB:9956062
C:Superfamily: CARS cyclophilin; cyclophilin homology
F:7-177/Domain: cyclophilin homology <CYP>
Alignment Scores:
Pred. No.: 0.00753 Length: 754
Score: 129.00 Matches: 76
Percent Similarity: 38.48% Conservative: 51
Best Local Similarity: 23.03% Mismatches: 145
Query Match: 6.57% Indels: 58
Gaps: 12
US-10-046-935-2234 (1-1116) x JC5314 (1-754)
QY 4 GACGCTCGCGCGTCCGAGAAAGATCTCAGAGTAAAGAGAACTTAAAGAAATTCAGA 63
Db 437 AspIleArgArgAsnSerGluLysAspAspLysTyrLysAsnLysValLysLysArgAla 456

Db 258 uAsnAspGluProLeuLeuLysLysAlaAlaSerAlaLysAlaValGlnThrAspLysPr 278
 QY 634 -----AAGACCGTGGATCGGTACATGAATGAA-----GATGACCT 668
 Db 278 oThrGlyGluHisValLysThrSerAspSerTrpLeuSerLysThrArgPheAspGlnPh 298
 QY 669 GCCAGAACCGCTCGCTCCAGATCCTCGT----- 699
 Db 298 eProLeuSerProLeuSerLeuLysAlaIleLysAspAlaGlyPheGluThrMetThrVa 318
 QY 700 -----ACCTCCGCATATAATTCGC 720
 Db 318 lValGlnGluAlaThrLeuProIleIleLeuGln 329
 RESULT 11
 T43213
 ENBP1 protein - barrel medic
 N:Alternate names: chloroplast-localised DNA-binding protein PD3 homolog
 C:Species: Medicago truncatula (barrel medic)
 C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
 C:Accession: T43213
 R:Christiansen, H.; Barker, D.G.
 A:Description: ENBP1 is homologous to a chloroplast-localised protein from pea, PD3.
 A:Reference number: Z22340
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1701 <CHR>
 A:Cross-references: UNIPROT:082022; EMBL:AJ002479; NID:el1317985; PID:el1317986; PIDN:CAAC
 A:Experimental source: cultivar Jemalong
 C:Genetics:
 A:Gene: ENBP1
 A:Introns: 856/3; 900/2; 933/3; 978/1; 1000/2; 1220/3; 1329/3; 1585/3; 1610/1; 1638/3; 1
 Alignment Scores:
 Pred. No.: 0.122 Length: 1701
 Score: 115.50 Matches: 70
 Percent Similarity: 32.77% Conservative: 47
 Best Local Similarity: 19.61% Mismatches: 129
 Query Match: 5.89% Indels: 111
 DB: 2 Gaps: 13
 US-10-046-935-2234 (1-1116) x T43213 (1-1701)
 QY 244 AGGGAGCAACCAACAAAAAGCAGAGTCCCGCCAGCCCTCAGAGAATTCGTGACTGAT 303
 Db 628 LysGlySerLysAsnLysLys----- 634
 QY 304 TCCAACTCCGATTCAGAGATGAAAGTGAATTTTTCGAGAAAAGG----- 354
 Db 635 LysAsnIleAlaGlyGluAspGlyAsnLysLeuHisLysGluLysLysArgGlyTrp 654
 QY 355 -----GCTTTAATATAAGCAACAAAGCAATCTTCGAAATCATGTCT 402
 Db 655 ProLysGlyPheCysLeuLysProLysGluIleAlaAlaArgLeuAspGluLysIleGlu 674
 QY 403 GAATTAGAAAGCTTCCTCGCTGGCTGGTGGAGACATCCCTCCAGGCTCCGACTCA 462
 Db 675 ArgArgGlyArgProLysGlySer-----GlyMetLysProLysGluThrAlaValGln 692
 QY 463 CAATCAGGAGACCGGAGCGGTACATTCGCGGTGTTGCTCCAGGAGA----- 513
 Db 693 LeuAspAlaLysIleGluArgGlyArgProLysGlyAlaGlyLysLysProLysGlu 712
 QY 514 -----AACCTGAACGAGAGCTCGTCTCTT----- 540
 Db 713 IleValValArgLeuAspThrLysIleGluArgArgGlyArgProLysGlySerGlyLys 732
 QY 541 -----ACAGGTCAAGGTCCTCGG 558
 Db 733 LysGlnLysGluValAlaSerGlnLeuAlaLeuGlnIleGluSerGlnLysSerThrArg 752

QY 559 ATCTCGGGTCCCTTCGACTCTA----- 582
 Db 753 ValAspGlyAlaLeuSerThrIleValProHisLysHisIleGlnGluSerIleSer 772
 QY 583 -----CCCATGGAGAGCCAGCTTTTGARKMTWMTACATGTTGGTGAGAAAG 630
 Db 773 ProLeuLysAspProValAsnLysGluGluLysSerAspPheValLeuGluCysSerLys 792
 QY 631 -----AGGAAGACCGTGGATGCTACATCAATGAAGATGACCTGCCAGAGC 678
 Db 793 AspSerGlyIleGluLysIleThrLysGlyLeuMetSerLysSerGlyAspValHisLys 812
 QY 679 CGTCGCTCCAGATCATCCGTGACCTTCGCCATATAAATTCGCCACGAGAGAAATTACA 738
 Db 813 ArgCysSerGluArgLeuArgThrLeuLeuThrAspHisLysAsnSerGlnAspValGlu 832
 QY 739 GAGGAGAGTTGGAGAACGCTCTGCAGCAATTCGAGAGAAAGATATATACCGTTCACGTG 798
 Db 833 ValGlu-----GluThrPheCysGluAsnGluValGluGluAlaIleAspHisGluLeu 850
 QY 799 GGCTCTACT----- 810
 Db 851 GluSerSerAspLeuMetGlyGluProGluThrLysLysGluProArgAsnLeuArgCys 870
 QY 811 CATCAATGCCGTGAGAACT-----ATTGATACCAAAACAACTGCAGAAACCCA 861
 Db 871 HisGlnCysTrpLysLysSerArgThrGlyIleValValCysThrLysCysLys----- 888
 QY 862 GACTGCTGGGGCTTCGAGGCCAGTTCGTGCCCCCTGCCTTCGAAACCGTTATGTTGAA 921
 Db 889 -----ArgLysLysTrpCysTrpGluCysIleAlaLysTrpTrpGlnAsp 903
 QY 922 GAGGTCAGGAGTCTCTGCTGATCCGAATCGCATTCGCCCTTGTGAGAACTCTGC 981
 Db 904 LysThrArgGluGluLe-----GluThrAlaCysProPheCysLeuAspTyrCys 920
 QY 982 AACTGCAGTTTCTGCCGAGCAGATGAGACGGTGTGAGACTGGGGTCTCTGTGTTATTTA 1041
 Db 921 AsnCysArgLeuCysLeuLysLysThr----- 930
 QY 1042 GCCAAATATCATGGCTTTGGGAATGTGCATGCTACTTCAAAGAGCTTGAAA 1092
 Db 931 SerThrMetAsnGlyAsnGlyGluAlaAspValLysLeuGlnLys 947
 RESULT 12
 G59436
 KIAA1304 protein [imported] - human
 C:Species: Homo sapiens (man)
 C>Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 09-Jul-2004
 C:Accession: G59436; H59436
 R:Nagase, T.; Kikuno, R.; Ishikawa, K.I.; Hirose, M.; Ohara, O.
 DNA Res. 7, 65-73, 2000
 A:Title: Prediction of the coding sequences of unidentified human genes. XVI. The complete
 A:Reference number: G59436
 A:Accession: G59436
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1051 <NAG>
 A:Cross-references: UNIPROT:Q7Z6B7; GB:BAA92542; PID:g7242963; PIDN:BAA92542.1
 R:Ohara, O.; Nagase, T.; Kikuno, R.
 submitted to GenBank, January 2000
 A:Reference number: H59436
 A:Accession: H59436
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1051 <OHA>
 A:Cross-references: GB:BAA92542; PID:g7242963; PIDN:BAA92542.1
 Alignment Scores:
 Pred. No.: 0.127 Length: 1051
 Score: 115.00 Matches: 77
 Percent Similarity: 31.70% Conservative: 46
 Best Local Similarity: 19.85% Mismatches: 129

Db 318 LysLysAlaGlnGluAsnAspValLysLysGlnLeuThrLysLysSerMetProAla 337
QY 565 GGG-----TCCCTTGAGCTCTACCCCATGGRGAGAGCCAGT 600
Db 338 GlyGluArgTyrSerGlnGluSerSerGlyLeuAspArgProLeuLysGlnLys 357
QY 601 TTKARKMTMMGTACATGTTG-----GTGAGAAAGAGAGACC 639
Db 358 LeuAspGlySerValThrValArgAspGlyTyrAspThrThrIleLeuGlnAsnIleThr 377
QY 640 GTGGATGCTACATGAATGAACATGACCTGCCAGAGCCGTCCTCAGATCATCCGTTG 699
Db 378 SerAspGly-----LysLysAspThrGlySerPheLysArgProArgAspLysVal 394
QY 700 ACCCTTCCGCATATAATTCCGCCAGTGAAGAAATATACAGAGCAGGAGTGGAGAACGTC 759
Db 395 Thr-----TleGluGluValProGluLysArgSerPheVal 407
QY 760 TGCAGCAATTCTCGAGAGAGATATATAACCGTTCACTGGGCTCTACTTGTTCATCAATGC 819
Db 408 -----LysAsnArgAspLeuValSerValSerGluGlyLysThrThrLysThrVal 425
QY 820 CGTCAGAGACTATTGATACCAACAAACAC-----849
Db 426 ThrGluLysGlyLysSerLysLysProSerPheGlyArgAlaGluAspLysMetSerAla 445
QY 850 -----TGAGAAACCCAGACTGCTGGGCG 873
Db 446 AspAspAsnGluArgAsnTyrGlnValThrGluValCysArgArgProAspAlaArgVal 465
QY 874 GTT-----CGAGGCGAGTCTGTGCGCCCTGCTCGAAACCGTTATGTT 918
Db 466 ValLysSerLeuAspAlaArgMetTyrPheCysLysAsnThrIle-----480
QY 919 GAAGAGTCAGGATGCTGCTG-----GATCCGAACTGCATTCGCCGCTTGTGCA 972
Db 481 -----LeuLysGlyAlaValLeuPheTyrSerProSer-----HisCysMet-----494
QY 973 GGAATCTGCACTGCAGTTCTTCTGCCGCGAGATGGACGGTGTGCGACTGGGGTCTTT 1032
Db 495 ---LeuTyrSerCysSer-----AspValLeu 502
QY 1033 GGTATATTAGCAAAATATCAT-----GGCTTTGGAAATGTGATCGCTACTTGA 1083
Db 503 SerPheSerSerTyrTyrValThrThrMetLeuGlyAsnIleGluProTyrGluGlu 522
QY 1084 AGCCTGAAACAG 1095
Db 523 SerMetArgGlu 526

RESULT 14
T10955
early nodulin binding protein 1 - spring vetch
C:Species: Vicia sativa (spring vetch, tare)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T10955
R:Christiansen, A.; Hansen, A.C.; Vijn, I.; Pallisgaard, N.; Larsen, K.; Yang, W.C.; Biss
submitted to the EMBL Data Library, December 1995
A:Description: A novel type of DNA binding protein interacts with a conserved sequence i
A:Reference number: Z17228
A:Accession: T10955
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1641 <CHR>
A:Cross-references: UNIPROT:Q41700; EMBL:X95995; NID:gl360633; PID:e225826
C:Genetics:
A:Note: ENBP1
C:Keywords: DNA binding

Alignment Scores:
Pred. No.: 0.199 Length: 1641
Score: 113.00 Matches: 73
Percent Similarity: 32.50% Conservative: 44

Best Local Similarity: 20.28% Mismatches: 125
Query Match: 5.76% Indels: 118
DB: 2 Gaps: 17
US-10-046-935-2234 (1-1116) x T10955 (1-1641)
QY 67 GTGAAGTTGATTTCCATGGAACCTCGTCATCTCTGATGACAGTTGTGACAGCTTGTCT 126
Db 585 lleLysLeuIleGlyMetGluSer-----GluAlaAlaValSerValGly 600
QY 127 TCTGATATATTTTGAACACAGAGGCTGCAGTCAGTTCGGGAAGGCTGTAGGACCCGACG 186
Db 601 AspGlnGluLeuValValGlnProLeuValLysValArg-----613
QY 187 CAGTGCAGCACTCTCGACCTCTCAGGCTGGCGATGAAGTTTCCAGCGGAGTACCAGG 246
Db 614 -----PheArgMetLeuAsnProLysMetGlyArgPro-----Lys 625
QY 247 GGAGCAACCAACAAAGACAGAGTCCGCCAGCCCTCAGAGAATTTCTGTGATGATCC 306
Db 626 GlySerLysAsnLys-----Lys 632
QY 307 AACTCCGATTCAAGATGAAGTGAATGAATTTTGGAGAAAG-----354
Db 633 AsnValAspGlyGluAlaGluAsnGlyLeuHisLysGluGlyLysArgGlyArgPro 652
QY 355 -----GCTTTAAATATAAGCAA-----AACAAAGCAATGCTTCAAAACTCATG 399
Db 653 LysGlySerValAsnAsnProLysGluThrGlyAsnGluLysIleAlaLysGlyLeuVal 672
QY 400 TCTGAATTAGAAAGCTTCCCTGGCTGCTTCCGTGGAAGACATCCCTCCAGGCTCCGAC 459
Db 673 SerGlu-----Ser 675
QY 460 TCACATCAAGAGACCGGAGCGGTACATTCGCCGGTGTGTCTTCCAGGAAACCT 519
Db 676 SerAsnValHisLysIleGluArgGlyArgPro-----LysGlySerAlaProAsnPro 694
QY 520 GAACGGAGAGCTGCTCTT-----ACCAGGTCAAGGTCCCGGATCTCTC 564
Db 695 LysGluAsnAlaSerArgLeuAspAlaGluLeuGluArgGluLysAsnThrHisValTyr 714
QY 565 GGGTCCCTTGAC---GCTCTACCATGGAGAGCCAGTTTKARKMTMMGTACATGTTG 621
Db 715 GlyIleLeuSerThrThrMetProHisLysHisIleHisGluGluSerIleLeuLeu 734
QY 622 -----GTCAGAAAGAGGAGACCGTGGAT-----645
Db 735 GluAspHisValAsnLysLysAspAlaAspPheValLeuGluCysSerLysGluSer 754
QY 646 -----GGCTACATGAATGAAGATGACCTGCCAGAGCCGT---681
Db 755 GlyIleGluLysIleAlaLysGlyLeuValSerGluSerAspAsnValHisLysThrHis 774
QY 682 -----CGCTCCAGATCATCGTACCGTCCCGATATA 714
Db 775 AspValGluValGlyAspIlePheTyrGluLysGluValLysGluThrIleAspHisArg 794
QY 715 ATTCCGCCAGTGAAGAATACAGAGGAGGTTGGAGACGCTCTGAGCAATCTCGA 774
Db 795 LeuGluProSerAspMetGlyAspCysGluThrLysLysGluProArgAsnSerArg 814
QY 775 GAGAAGATATATAACCGTTCACTGGGCTCTACTTGTTCATCAATCGCTCAGAGACT---831
Db 815 -----CysHisGlnCysTriPylsLysSerArg 823
QY 832 -----ATTGATACCAAAACAACTGCAGAAACCCAGACTGCTGGGGGTTCGAGCCAG 885
Db 824 ThrGlyLeuValValCysSerLysCysLys-----LysLysLys 836
QY 886 TTCTGTGGCCCTGCTTCGAAACCGTTATGTTAGAGAGTCCAGGGATGCTCTGCTGGAT 945
Db 837 TyrCysTyrGluCysValAlaLysTyrPyrGlnAspLysThrArgGluGluLeu-----854

QY 946 CGAACTGGCATTGCCCGCTTGTGCGAGGATCTCGAACTGTCAGTTTCTGCGGCGAGCGA 1005
 Db 855 ---GluThrAlaCysProPheCysLeuAspTyrCysAsnCysArgMetCysLeuLysLys 873

RESULT 15
 A47392
 chromodomain-helicase-DNA-binding protein, CHD-1 - mouse
 N:Alternate names: KYP protein
 C:Species: Mus musculus (house mouse)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Dec-2002
 C:Accession: A47392; S21568
 R:Delmas, V.; Stokes, D.G.; Perry, R.P.
 Proc. Natl. Acad. Sci. U.S.A. 90, 2414-2418, 1993
 A:Title: A mammalian DNA-binding protein that contains a chromodomain and an SNF2/SWI2-1
 A:Reference number: A47392; MUID:93211972; PMID:8460153
 A:Accession: A47392
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1711
 A:Experimental source: S194 plasmacytoma cells
 A>Note: sequence inconsistent with the nucleotide translation
 A>Note: sequence extracted from NCBI backbone (NCBIN:128272, NCBIPI:128273)
 R:Delmas, V.; Perry, R.P.
 submitted to the EMBL Data Library, May 1992
 A:Description: KYPB, a mammalian protein that contains the SNF2/SWI2 helicase domain ale
 A:Reference number: S21568
 A:Accession: S21568
 A:Molecule type: mRNA
 A:Residues: 772-1711 <DE2>
 A:Cross-references: EMBL:X66028
 C:Superfamily: chromodomain helicase CHD1; chromobox homology
 C:Keywords: DNA binding
 F:293-336/Domain: chromobox homology <CB1>
 F:387-427/Domain: chromobox homology <CB2>

Alignment Scores:
 Pred. No.: 0.27 Length: 1711
 Score: 111.50 Matches: 82
 Percent Similarity: 38.04% Conservative: 58
 Best Local Similarity: 22.28% Indels: 95
 Query Match: 5.68% Gaps: 18
 DB: 1

US-10-046-935-2234 (1-1116) x A47392 (1-1711)

QY 79 TCATGGAACCTCGTCATCTCTGATGACAGTGTGACAGCTTGTCTCTGATATATTT 138
 Db 39 SerSerAspGlySerSerSerGlnSerGlySerSerAsp-----SerAspSerGly 55

QY 139 GCAACACGAGGCTGCAGTCAGTTCGGGAGGCTGTAGGACCGGACGAGTCGAGGCAC 198
 Db 56 SerAspSerGlySerGlnSerGluSerGluSerAspThrSerArgGluAsnLysValGln 75

QY 199 TCTGACCTCTCAGGTG---GCGATGAAGTTTCCAGCGCGAGT-----240
 Db 76 AlaLysProProLysValAspGlyAlaGluPheTrpLysSerSerProSerIleLeuAla 95

QY 241 ---ACCAGGGAGACCAACAAACAAA-----GCAAGTCCCGCCAGCCC---282
 Db 96 ValGlnArgSerAlaMetLeuArgLysGlnProGlnGlnAlaGlnGlnArgProAla 115

QY 283 TCAGAGAATCTGTGATCTCACTCCGATTCGATTCAGAGATCAAGTGAATGATTTT 342
 Db 116 SerSerAsnSerGlySerGluGluAspSerSerSerSerGluAspSer-----131

QY 343 TTGAGAAAAAGGCTTTAAATATAAAGCAAAAGCAATGCTTGCAAAACCTATGCT 402
 Db 132 ---AspAspSerSerSerGlyAlaLysArgLysLys-----His 143

QY 403 GAATTAGAAAGCTTCCTGGCTGCTTCGTTGGGAGAGATCCCTCCAGGCTCCGACTCA 462
 Db 144 AsnAspGluAspTrpGlnMetSerGlySerGlySerProSerGlnLeuGlySerAspSer 163

Search completed: October 12, 2004, 15:19:50

Job time : 43 secs

QY 463 CAATCAAGGAGACCGCGA---AGCGGTACATTCCCGGGTGTGCTTCCAGGAGAAACCCCT 519
 Db 164 GluSerGluGluGluArgAspLysSerSerCysAspGlyThrGluSerAspTyrGluPro 183

QY 520 GAAACGAGAGCTCGT-----CCTCTTACAGGTCAAGTCC-----CGG 558
 Db 184 LysAsnLysValArgSerArgLysProGlnAsnArgSerLysSerLysAsnGlyLysLys 203

QY 559 ATCTCTCGGGTCCCTTGACGCTCTACCCATGGGAGAGCCAGTTTGAKRMWMTACATG 618
 Db 204 IleLeuGly-----206

QY 619 TTGGTGAAGAAAGAGACCGTGGATGGCTACATGAATGAAGATGACCTG-----669
 Db 207 -----GlnLysLysArgGlnIleAspSerSerGluAspGluAspGluAspTyrAsp 224

QY 670 -----CCCAGAAGCCGTGCTCCAGATCATCCGTGACCCCTTCCGCATATATAATTCGCCCA 723
 Db 225 AsnAspLysArgSerSerArgArgGlnAlaThrValAsnValSerTyrLysGluAspGlu 244

QY 724 GTGGAAGAAATTCACAGAGAGAGTGGAGAACGTCTGCAGCAATCTCGAGAGAAGATA 783
 Db 245 GluMetLysThrAspSerAspLeuLeuGluValCysGlyGluAspValProGlnPro 264

QY 784 TATAACCGTTCACCTGGCTCTACTTGTTCATCAATCCGTCGAGAGACTATTGATACCAA 843
 Db 265 GluAspGluGluPhe-----GluThrIleGluArgVal 275

QY 844 ACAAACTGCAGAAACCCAGACTGTGGGGCGTTCGAGGCCAGTTCTGTGGCCCTGCTT 903
 Db 276 MetAspCysArg-----ValGlyArgLysGlyAlaThrGlyAlaThr 289

QY 904 CGAAACCGTTCGTGAAGAGTCAAGGATGCTCTGTGGATCCGAACTGGCATTCGCCG 963
 Db 290 ThrThrIleTyrAlaValGluAlaAspGly-----AspPro-----301

QY 964 CCTTCTCGAGGAATCTGCAACTGCAGTTTCTGCCGCGACGAGATGGACGGTGTGCGACT 1023
 Db 302 -----AsnAlaGlyPheGluArgAsnLysGlu-----Pro 311

QY 1024 GGGGTCTTGTGTATTTAGCCAAATATCATGCTTTGGGAATGTGCAT-----GCC 1074
 Db 312 GlyAspIleGlnTyrLeuIleLysTrpLysGlyTyrSerHisIleHisAsnThrTrpGlu 331

QY 1075 TACTTGAAGAGCTGAAACAGGAA 1098
 Db 332 ThrGluGluThrLeuLysGlnGln 339

Blank

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 12, 2004, 15:05:41 ; Search time 171.5 Seconds
(without alignments)
7488.258 Million cell updates/sec

Title: US-10-046-935-2234
Perfect score: 1962
Sequence: 1 atggacgctgcgcggtgcc.....aattgaaatgcaagataa 1116

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 3650362

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-O=/cn2_1/USPTO spoop_p/US10046935/runat_12102004_155225_15243/app_query.fasta_1.1287
-DB=Uniprot_02 -QFMT=fastan -SUFFIX=rxp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosun62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10046935 @CN1_1.183 @runat_12102004_155225_15243 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPHLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Uniprot 02.*

- 1: uniprot_sprot.*
- 2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1929	98.3	371	Q9BWT1	Q9bwt1 homo sapien
2	1879	95.8	450	Q9GSP5	Q9gsp5 homo sapien
3	1612.5	82.2	382	Q9D0M2	Q9d0m2 m mus muscu
4	1609.5	82.0	382	Q6NZE5	Q6nze5 mus musculu
5	1609.5	82.0	382	Aah66169	Aah66169 mus muscu
6	1607.5	81.9	382	Q8C1A0	Q8c1a0 mus musculu
7	1155	58.9	220	Q96BV8	Q96bv8 homo sapien
8	786	40.1	207	Q8NBY9	Q8nby9 homo sapien
9	752	38.3	442	Q96C70	Q96c70 homo sapien
10	752	38.3	453	Q86YT0	Q86yt0 homo sapien
11	752	38.3	453	Aah25242	Aah25242 homo sapi
12	752	38.3	454	Q96GN5	Q96gn5 homo sapien
13	743	37.9	453	Q6P1L4	Q6p1l4 homo sapien
14	743	37.9	453	Aah32576	Aah32576 homo sapi
15	736	37.5	438	Q922M5	Q922m5 mus musculu
16	602.5	30.7	202	Q9NPV2	Q9npv2 homo sapien

17	289.5	14.8	462	2	Q23175	Q23175 arabidopsis
18	250.5	12.8	515	2	Q9FXE1	Q9fxe1 arabidopsis
19	233.5	11.9	544	2	Q9FEV3	Q9ffv3 arabidopsis
20	233.5	11.9	572	2	Q8RW95	Q8rw95 arabidopsis
21	225.5	11.5	495	2	Q9FYG6	Q9fyg6 arabidopsis
22	225.5	11.5	297	2	Q9AURO	Q9aur0 oryza sativ
23	216.5	11.0	297	2	Q9AURO	Q9aur0 oryza sativ
24	171.5	8.7	268	2	Q8H9A2	Q8ha2 homo sapien
25	170	8.7	555	2	Q80469	Q80469 arabidopsis
26	146	7.4	851	2	Q6S384	Q6s384 arabidopsis
27	142	7.2	950	2	Q04024	Q04024 arabidopsis
28	132.5	6.8	566	2	Q6Q150	Q6q150 arabidopsis
29	132.5	6.8	566	2	AAS75310	Aas75310 arabidops
30	131.5	6.7	906	2	Q9LV75	Q9lv75 arabidopsis
31	129	6.6	754	1	Q48794	Q48794 arabidopsis
32	124.5	6.3	1007	1	PR4G_HUMAN	Q13427 homo sapien
33	124.5	6.3	1007	2	Q8BND8	Q8bnd8 mus musculu
34	123.5	6.3	1458	2	Q9HE72	Q9he72 neurospora
35	122	6.2	820	2	Q60585	Q60585 homo sapien
36	121.5	6.2	497	2	Q8C5G1	Q8c5g1 mus musculu
37	119.5	6.1	1150	2	Q9W0H4	Q9w0h4 drosophila
38	118.5	6.0	822	2	Q6DFC7	Q6dfc7 xenopus lae
39	118.5	6.0	879	2	Q80Y24	Q80y24 mus musculu
40	118	6.0	2759	2	Q9VID9	Q9vid9 drosophila
41	117.5	6.0	595	2	Q9Z8Z9	Q9zrz9 arabidopsis
42	117.5	6.0	827	2	Q9FNM7	Q9fmn7 arabidopsis
43	117	6.0	1001	2	Q75DD5	Q75dd5 ashbya goss
44	117	6.0	1001	2	AAS50862	Aas50862 ashbya go
45	116.5	5.9	514	2	Q9NWD4	Q9nwd4 homo sapien

ALIGNMENTS

RESULT 1

ID	Q9BWT1	PRELIMINARY;	PRT;	371 AA.
AC	Q9BWT1;			
DT	01-JUN-2001 (TrEMBLrel. 17, Created)			
DT	01-JUN-2001 (TrEMBLrel. 28, Last sequence update)			
DT	01-OCT-2004 (TrEMBLrel. 28, Last annotation update)			
DE	C-Myc target JP01 (Cell division cycle associated protein 7, isoform 2) (Hypothetical protein FLJ14736).			
GN	Name=JP01; Synonyms=CDCA7;			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21611205; PubMed=11598121;			
RA	Prescott J.E., Osthus R.C., Lee L.A., Lewis B.C., Shim H.,			
RA	Barrett J.F., Guo Q., Hawkins A.L., Griffin C.A., Dang C.V.;			
RT	"A novel c-Myc-responsive gene, JP01, participates in neoplastic transformation."			
RL	J. Biol. Chem. 276:48276-48284 (2001).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Pancreas;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Locquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISUE=Pancreas;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
EX PubMed:14702039;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto K., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Mubashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiya M., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yanazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Ohtani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs";
RL Nat. Genet. 36:40-45 (2004).
DR EMBL; AY029179; AK31591.1; -;
DR EMBL; BC027966; AAH27966.1; -;
DR EMBL; AK027642; BAB55258.1; -;
KW Cell division.
SQ SEQUENCE 371 AA; 42573 MW; 30A244E3057D9C43 CRC64;

Alignment Scores:

Pred. No.: 4,6e-153 Length: 371
Score: 1929.00 Matches: 364
Percent Similarity: 98.11% Conservative: 0
Best Local Similarity: 98.11% Mismatches: 7
Query Match: 98.32% Indels: 0
DB: Gaps: 0

US-10-046-935-2234 (1-1116) x Q9BWT1 (1-371)

QY 1 ATGGAGCGCTCGCGCGTCCGCGGAGAGATCTCAGAGTAAAGAACTTAAAGAAATTC 60
DB 1 MetAspAlaArgArgValProGlnLysAspLeuArgValLysLysAsnLeuLysPhe 20
QY 61 AGATATGTGAAGTTGATTTCCATGGAAACCTCGTATCTCGATGACAGTTGTGACAGC 120
DB 21 ArgTyrValLysLeuIleSerMetGluThrSerSerSerSerSerSerSerSerSer 40
QY 121 TTGTGCTTCTGATAATTTTGCACACAGAGCTCGAGTTCGGGAAGCGCTGTAGAGC 180
DB 41 PheAlaSerAspAsnPheAlaAsnThrArgueGlnSerValArgIleGlyCysArgThr 60

QY 181 CGCAGCCAGTCCAGGCACTCTGGACCTCTCAGGCTGGCGATGAAGTTTCCAGCGCGGAGT 240
DB 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80
QY 241 ACCAGGGAGCAACCAACAAAAAGACAGAGTCCCGCCAGCCCTCAGAAATTCGTGACT 300
DB 81 ThrArgGlyAlaThrAsnLysAlaGluSerArgGlnProSerGluAsnSerValThr 100
QY 301 GATTCCAACTCCGATTTCAGAAAGTGAAGTGAATTTTTCGAGAAAGGGCTTTA 360
DB 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 120
QY 361 AATATAAGCAAAACAAAGCAATCTTCCAAAACCTCATGTCTGAATTAGAAGCTTCCT 420
DB 121 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140
QY 421 GGCTCGTTCCGTGGAGACATCCCTCCAGGCTCCGACTCACATCAAGAGACGGCA 480
DB 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160
QY 481 AGGCTCATCTCCCGGGTGTCTTCCAGGAGAAACCCCTGAACGAGAGCTCGTCTCTT 540
DB 161 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgAlaArgProLeu 180
QY 541 ACCAGTCAAGTCCCGGATCTCGGTCCTTGAAGCTCTACCATCGRGAGAGCCAGT 600
DB 181 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGlu 200
QY 601 TTGARKMTWGTACATCTTGGTGAAGAGGAGAGACCGTGATGGCTACATGAATGAA 660
DB 201 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220
QY 661 GATGACCTGCCAGAAAGCGTCCGCTCCAGATCATCTCCGACCTTCCGATATAATTCG 720
DB 221 AspAspLeuProArgSerArgArgSerArgSerValThrLeuProHisIleArg 240
QY 721 CCAGTGAAGAAATTCAGAGAGAGAGTGGAGAACCTGTGAGCAATTCGAGAGAGAG 780
DB 241 ProValGluGluIleThrGluGluLeuGluAsnValCysSerAsnSerArgGluLys 260
QY 781 ATATATACCGTTCACCTGGCTCTACTTGTCTCATCAATCCGTCAGAAAGACTATTGATCC 840
DB 261 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 280
QY 841 AAAACAAACTGCAGAAACCCAGCTGCTGGGCGCTTCGAGGCGAGTCTGTGGCCCTGC 900
DB 281 LysThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCys 300
QY 901 CTTCCAAACCGTTATGGTGAAGAGTCCAGGATGCTCTGCTGGATCCGAACTGGCATTCG 960
DB 301 LeuArgAsnArgTyrGlyGluGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 320
QY 961 CGGCTTGTCCAGGAATCTGCAACTGCAGTTCCTCCGCGCAGCAGATCGAGCTGTGCG 1020
DB 321 ProProCysArgGlyLysCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 340
QY 1021 ACTGGGCTCCTTGTGTATTAGCCAAATATCATGGCTTGGGAATGTGCATGCTCTTGG 1080
DB 341 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 360
QY 1081 AAAAGCTGAAACAGAAATTTGAAATGCAAGCA 1113
DB 361 LysSerLeuLysGlnGluPheGluMetGlnAla 371
RESULT 2
Q9BSP5 ID Q96SP5 PRELIMINARY; PRT; 450 AA.
AC Q96SP5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ14722.
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RX PubMed=14702039;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
 RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
 RA Sudo H., Hosoliri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
 RA Abe K., Kamihara K., Katsura N., Sato K., Tanikawa M., Yamazaki M.,
 RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
 RA Tanai H., Kimata M., Watanabe M., Hirao K., Chiba Y., Ishida S.,
 RA Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hotuta T., Kusano J.,
 RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
 RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
 RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
 RA Moriya S., Momiya H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogano S.,
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs".
 RL Nat. Genet. 36:40-45 (2004).
 DR EMBL; AK027628; BAB55245.1; -;
 DR Genbank; HGNC:14628; CDCA7.
 SQ SEQUENCE 450 AA; 51417 MW; D86A79E7981B773D CRC64;

Alignment Scores:

Pred. No.: 6.93e-149 Length: 450
 Score: 1879.50 Matches: 364
 Percent Similarity: 80.89% Conservative: 0
 Best Local Similarity: 80.89% Mismatches: 7
 Query Match: 95.80% Indels: 79
 DB: 2 Gaps: 1

US-10-046-935-2234 (1-1116) x Q96SP5 (1-450)

QY 1 ATGGACGCTGCGCGTCCGAGAGAGTCTCAGAGTAAGAGACTTAAGAAATTC 60
 Db 1 MetAspAlaArgValProGlnLysAspLeuArgValLysLysAsnLeuLysPhe 20
 QY 61 AGATATGCAAGTTGATTTCCTGGAACCTGTCATCTCTGATCAGAGTTGTGACAGC 120
 Db 21 ArgTyrValLysLeuLeuSerMetGluThrSerSerSerSerSerSerSerSer 40
 QY 121 TTTCCTCTCATAATTTTGCACACG----- 147
 Db 41 PheAlaSerAspAsnPheAlaAsnThrLysProLysPheArgSerAspIleSerGluGlu 60
 QY 147 ----- 147
 Db 61 LeuAlaSerValPheTyrGluAspSerAspAsnGluSerPheCysGlyPheSerGluSer 80
 QY 147 ----- 147
 Db 81 GluValGlnAspValLeuLeuAspHisCysGlyPheLeuGlnLysProArgProAspValThr 100
 QY 147 ----- 147

Db 101 AsnGluLeuAlaGlyIlePheHisAlaAspSerAspGluSerPheCysGlyPheSer 120
 QY 148 -----AGGCTGCAGTCTCAGTTCGGGAAGGCTGTAGGACCCGC 183
 Db 121 GluSerGluIleGlnAspGlyMetArgLeuGlnSerValArgGluGlyCysArgThrArg 140
 QY 184 AGCCAGTGCAGGCACCTCTGAGGCTGGCGATGAAGTTTCCAGCGCGAGTAGTACC 243
 Db 141 SerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSerThr 160
 QY 244 AGGGAGCAACCAACAAACAAAGACAGAGTCCCGCCAGCCCTCAGAGAAATCTGTGACTGAT 303
 Db 161 ArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThrAsp 180
 QY 304 TCCAACTCCGATTCAGAGATGCAAGTGAATTTTGGAGAAAGGGCTTTAAAT 363
 Db 181 SerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeuAsn 200
 QY 364 ATAAAGCAAAACAAACGAATCTTGCACAACTCATGTCTGAATTTAGAAGCTTCCTCGC 423
 Db 201 IleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPheProGly 220
 QY 424 TCGTTCCTGGAGACATCCCTCCAGCTCCGACTCACATCAAGAGACCGGAGG 483
 Db 221 SerPheArgGlyArgHisProLeuProLysSerAspSerGlnSerArgProArgArg 240
 QY 484 CGTCAATTCCTGGGCTGCTTCCAGGAGAACCTCGAACGAGAGCTCGTCTCTTACC 543
 Db 241 ArgThrPheProGlyValAlaSerArgArgAsnProGluArgAlaArgProLeuThr 260
 QY 544 AGGTCAAGGTCTCCGGATCTCGGGTCCCTTGACGCTCTACCCATCGRAGAGCCAGTTTK 603
 Db 261 ArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGluGlu 280
 QY 604 GARKMTMGTTACATCTTGGTGAAGAGAGACCGTGGATCGGTACATGATGAAGT 663
 Db 281 GluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGluAsp 300
 QY 664 GACCTGCCAGAGCGCTCGCTCCAGATCATCCGTGACCTTCCGATATAATTCGCCCA 723
 Db 301 AspLeuProArgSerArgArgSerArgSerValThrLeuProHisIleIleArgPro 320
 QY 724 GTGGAAGAAATTACAGAGAGAGTGGAGAACCGTCTGCAGCAATTCGAGAGAGATA 783
 Db 321 ValGluGluIleThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLysIle 340
 QY 784 TATAACCGTTTCACTGGGCTCTACTTGTCTCATCAATGCCGTGAGAGACTATTGATACCAA 843
 Db 341 TyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThrLys 360
 QY 844 ACAAACTGCAGAAACCCAGACTGCTGGGCGTTCGAGGCGCAGTCTGTGGCCCTGCCCTT 903
 Db 361 ThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCysLeu 380
 QY 904 CGAAACCGTTATGGTGAAGAGGTGAGGATGCTCTGTGGATCCGAACTGGCATTGCCCG 963
 Db 381 ArgAsnArgTyrGlyGluValArgAspAlaLeuLeuAspProAsnTrpHisCysPro 400
 QY 964 CTTGTGAGGAATCTGCAACTGCACTGAGTTTCTCCGCGCAGCAGATGCGGCTGGGACT 1023
 Db 401 ProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAlaThr 420
 QY 1024 GGGGCTCTGTGCTATTAGCCAAATATCATGGCTTTGGGAATGTGCATCCCTACTTGA 1083
 Db 421 GlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeuLys 440
 QY 1084 AGCCTGAACACAGGAATTTGAAATGCAAGCA 1113
 Db 441 SerLeuLysGlnGluPheGluMetGlnAla 450

RESULT 3
 Q9DOM2
 ID Q9DOM2 PRELIMINARY; PRT; 382 AA.

Q9D0M2;
 01-JUN-2001 (TREMBlrel. 17, Created)
 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE Mus musculus 10 days embryo whole body cDNA, RIKEN full-length
 DE enriched library, clone:261002110 product:similar to C-MYC TARGET
 DE JPO1 (CDNA FJ174736 FIS, CLONE NT2RP3002181) (Mus musculus adult male
 DE tongue cDNA, RIKEN full-length enriched library, clone:2310021G01
 DE product:similar to C-MYC TARGET JPO1).
 GN Name=Cdc47;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Tongue, and Whole body;
 RX MEDLINE=99279253; PubMed=10349636;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Tongue, and Whole body;
 RX MEDLINE=21085660; PubMed=11217851;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Tongue, and Whole body;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Tongue, and Whole body;
 RX MEDLINE=20499374; PubMed=11042159;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Tongue, and Whole body;
 RX MEDLINE=20530913; PubMed=11076861;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuina T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RX MEDLINE=20530913; PubMed=11076861;
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
 RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
 RA Muramatsu M., Hayashizaki Y.;

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Tongue;
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Aono H., Arai A.,
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
 RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
 RA Tejima Y., Toya T., Yamamura T., Yamanaka I., Yasunishi A.,
 RA Yoshida K., Yoshino M., Muramatsu M., Hayashizaki Y.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK011289; BAB27519.1; -;
 DR EMBL; AK075882; BAC36027.1; -;
 DR MGD; MGI:1914203; Cdc47.
 SQ SEQUENCE 382 AA; 43837 MW; 51920663588CDD40 CRC64;
 Alignment Scores:
 Pred. No.: 1.76e-126 Length: 382
 Score: 1612.50 Matches: 309
 Percent Similarity: 86.72% Conservative: 24
 Best Local Similarity: 80.47% Mismatches: 36
 Query Match: 82.19% Indels: 15
 Gaps: 2
 DB:
 US-10-046-935-2234 (1-1116) x Q9D0M2 (1-382)
 QY 1 ATGACGCTCGCGTCCGAGAAAGATCTCAGAGTAAAGAAAGAAATTC 60
 DB 1 MetGluAlaArgAlaArgGlnLysAlaLeuLysVal---LysAsnLeuLysAspVal 19
 QY 61 AGATATGTGAAGTTGATTTCCATGGAACCTCGTCATCTCTGTGATGACAGTTGTGACAGC 120
 DB 20 ArgTyrMetLysLeuLysSerMetGluThrSerSerSerSerSerSerSerSerSerSer 39
 QY 121 TTTCTCTGTGTAATTTTTCGAAACACGAGGCTGAGTGTGCGGAAGGCTGTAGGACC 180
 DB 40 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnLeuAsnArgGluGlyCysArgThr 59
 QY 181 CGCAGCAGTGCAGGCACTCTGGACCTCTCAGGTGGCGATGAAGTTTCCAGCGCGAGT 240
 DB 60 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgAsn 79
 QY 241 ACCAGGGAGCAACCAACAAAGAGCAGTCCCGCCAGCCCTCAGAGAAATTCGTGACT 300
 DB 80 ThrArgArgAlaAlaSerLysLysAlaAlaProLysProSerGluSerSerSerSerSer 99
 QY 301 GATTCCCACTCCGATTCA-----GAGATGAAGTGGATGAT 339
 DB 100 AspSerHisSerAspSerGluGluGluGluGluGluGluGluGluGluGluGluMetAsn 119
 QY 340 TTTTGTGAGAAAGGGCTTTAAATATAAGCAAAAGCAATGCTTGCAGAACTCATG 399
 DB 120 PheLeuGluLysArgAlaLeuAsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMet 139
 QY 400 TCTGAATTAGAAAGCTTCCCTGGCTCGTTCGTGGAGAGACATCCCTCCAGGCTCCGAC 459
 DB 140 SerGluLeuGluSerPhePheProGlyLeuPheSerGlyArgHisSerLeuProGlyHisArg 159
 QY 460 TCACAA---TCAAGGAGACCGGAGCGTACATTCCTCCGGTGTGTCTCCAGGAGAAC 516
 DB 160 AlaLysAspLysSerProArgArgArgThrPheProGlyValAlaIleThrArgArgAsn 179
 QY 517 CCTGAACGAGAGCTCGTCTCTTTACAGGTCAGAGTCCCGGATCCCTCGGTCCTTCGAC 576
 DB 180 ProGluArgArgThrArgProLeuThrArgSerArgSerArgIleLeuGlySerLeuGly 199
 QY 577 GCTTACCCATG-----GRGAGAGCCAGTTTGTGARKMTWGTACATGTTG 621
 |||||

742	QY	GAGAGATTGGAGACGCTCTGCGCAATTTCTCGAGAGAGATATATAACCGTTCACCTGGCG	801
		:::	
259	Db	GluGluIleAArgAsnIleCysSerAsnSerArgGluIysIleTyAsnArgSerLeuGly	278
802	QY	TCTACTTGTCATCAATGCGGTTCAGAGACTATTGATACCAAAAACAACTCGAGAAACCCA	861
		:::	
279	Db	SerThrCysHisGlnCysArgGlnLysThrThrAspThrLysThrAsnCysArgAsnPro	298
862	QY	GACTGCTGGGGCGTTCTGAGGCCAGTTCTGTGCGCCCTGCCCTTCGAAACCGTTATGTGTGAA	921
		:::	
299	Db	AspCysTrpGlyIleAArgGlyGlnPheCysGlyProCysLeuAArgAsnArgTyGlyGlu	318
922	QY	GAGTCTCAGGATGCTCTGCTGGATCCGAACCTGGCATTCGCCCGCTTGTTCGAGGAATCTGC	981
		:::	
319	Db	GluValIysAspAlaLeuLeuAspProAsnTrpHisCysProCysArgGlyIleCys	338
982	QY	AACATGCAGTTTCTCCGCGCAGAGATGGACGGTGTGCGACTGGGGTCTTGTGTATTTA	1041
		:::	
339	Db	AsnCysSerPheCysArgsglnArgAspGlyArgCysAlaThrGlyValLeuValTyLeu	358
1042	QY	GCCAAATATCATGCTTTGGGAATGTGCATCCCTACTTTGAAAAGCCTGAAACAGGAATTT	1101
		:::	
359	Db	AlaArgTyHisGlyPheGlyAsnValHisAlaTyLeuLysSerLeuLysGlnLysPhe	378
1102	QY	GAATGCAAGCA 1113	
		:::	
379	Db	GluMetGlnAla 382	

RESULT 5

AAH66169 PRELIMINARY; PRT; 382 AA.

ID AAH66169; AC

DT 02-MAR-2004 (TrEMBürel. 27, Created)

DT 02-MAR-2004 (TrEMBürel. 27, Last sequence update)

DT 02-MAR-2004 (TrEMBürel. 27, Last annotation update)

DE Cell division cycle associated 7.

GN CDCA7.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OC NCBI_TaxId=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Unfertilized egg;

RA MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenman C.M., Schuler G.D.,

RA Altschul S.F., Ziegler B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieff F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Ustin T.B., Toshikiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Haie S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

RA Blakesley R.C., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Rodriguez A.C., Touchman J.W., Green E.D., Dickson M.C.,

RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RA "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RP
RC

SEQUENCE FROM N.A.
STRAIN=C57BL/6J; TISSUE=Unfertilized egg;

RA Strausberg R.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC066169; AAH66169.1; -.
KW Cell division.

```

SQ  SEQUENCE  382  AA;  43865  MW;  C8920663588CCFF7  CRC64;

Alignment Scores:
Pred. No.:      3.15e-126      Length:      382
Score:          1609.50        Matches:     308
Percent Similarity:  86.72%    Conservative:  25
Best Local Similarity:  80.21%  Mismatches:    36
Query Match:      82.03%      Indels:      15
DB:                2          Gaps:        5

```

US-10-046-935-2234 (1-1116) x AAH66169 (1-382)

QY	1	ATGACGGCTCGCGCGTCCCGCAGAAAGATCTCAGAGTAAAGAGAACTTAAAGAAATTC	60
Db	1	MetGluAlaArgAlaArgGlnLysAlaLeuLysVal---LysAsnLeuLysAspVal	19
QY	61	AGATATGTGAAGTTGATTTCATCGAAACCTCGTCATCTCTGATGACAGTTGTGCACAG	120
Db	20	ArgTyrMetLysLeuLysSerMetGluThrSerSerSerSerSerSerSerSerSerSerSer	39
QY	121	TTTGCTTCTGATAATTTTGGCAACACAGAGGCTGCAGTTCGCGGAAGCGCTGTAGGACC	180
Db	40	PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnLeuAsnArgGluGlyCysArgThr	59
QY	181	CGCAGCCAGTCAGCGACTCTGGGACTCTCAGGGTGGCGATGAAGTTTCCAGCGGACT	240
Db	60	ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMeLysPheProAlaArgAsn	79
QY	241	ACCAGGGGAGCAACCAACAAAAAAAGCAGAGTCCCGCCAGCCCTCAGAGAAATCTGTGACT	300
Db	80	ThrArgArgAlaAlaSerLysLysAlaAlaProProLysProSerGluSerSerAlaAsn	99
QY	301	GATTCCAACTCCGATTCA-----GAAGATGAAGTGAATGAAT	339
Db	100	AspSerHisSerAspSerGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu	119
QY	340	TTTTTGGAGAAAAGGGCTTTAAATATAAGCAAAACAAAGCAATCCTTGCAAACTCATG	399
Db	120	PheLeuGluLysArgAlaLeuAsnLleLysGlnAsnLysAlaMeLeuAlaLysLeuMet	139
QY	400	TCTGAATTAGAAAGCTTCCCTGGCTCGTTCGTTGGAAGACATCCCTCCAGGCTCCGAC	459
Db	140	SerGluLeuGluSerPheProGlyLeuPheSerGlyArgHisSerLeuProGlyHisArg	159
QY	460	TCACAA--TCAAGGAGACCGAGGCGTACATTCCGGGTGTGCTTCCAGGAGAAC	516
Db	160	AlaLysAspSerLysSerProArgArgArgThrPheProGlyValAlaThrArgArgAsn	179
QY	517	CCTGAACGGAGAGTCGTCTCTTACCAGGTCAAGGTCCTCGGATCCTCGGTCCTTGAC	576
Db	180	ProGluArgArgThrArgProLeuThrArgSerArgSerArgLleLeuGlySerLeuGly	199
QY	577	GCTCTACCCATG-----GRGAGAGCCAGTTTGAARKMTMWGTACATGTG	621
Db	200	AlaLeuProThrGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu	219
QY	622	GTGAGAAGAGGAGACCGTGGATGCTACATGAATGAAGATGACTCGCCACAGAGCCGT	681
Db	220	ValArgGlnArgLysSerMetAspSerTyrMetAsnAspAspValProArgSerArg	239
QY	682	CGCTCCAGATCATCCGTGACCTTCCGCATATAATTGCCAGTGGGAAGAAATTACAGAG	741
Db	240	ArgPro---GlySerMetThrLeuProHisIleLeuArgProValGluGluValThrGlu	258
QY	742	GAGGAGTTGGAGACGCTCTCAGCAATCTCAGAGAAAGATATATAACCGTTTCACTGGGC	801
Db	259	GluGluLeuArgAsnLleCysSerAsnSerArgGluLysLleTyrAsnArgSerLeuGly	278
QY	802	TCTACTTGTGATCAATCCGCTCAGAGACTATTGATACCAAAACMAACTGCAGAAACCCA	861
Db	279	SerThrCysHisGlnCysArgGlnLysThrThrAspThrLysThrAsnCysArgAsnPro	298
QY	862	GACTGCTGGGCGGTTTCGAGGCCAGTCTGTGGCCCTTCGCTTTCGAAAACGGTTATGGTGA	921

180 ProGluArgArgThrArgProLeuThrArgSerArgSerArgIleLeuGlySerLeuGly 199
QY 577 GCTTACCCATG-----GRGAGACCGAGTTCGARKMTWGTACATGTTG 621
Db 200 AlaLeuProThrGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 219
QY 622 GTGAGAAGAGAGAGAGCGTGGCTACATGATGAATGAATGATGCTCCCGAGAGCGGT 681
Db 220 ValArgGlnArgLysSerMetAspSerTyMetAsnAspAspValProArgSerArg 239
QY 682 CGCTCCAGATCATCGTGACCTCCGCATATAATATCCCGAGTGGAGAGAAATACAGAG 741
Db 240 ArgPro--GlySerMetThrLeuProHisIleArgProValGluGluValThrGlu 258
QY 742 GAGGAGTGGAGAGCGTCTGAGCAATTCGAGAGAGATATATAACCGTCTACTGGGC 801
Db 259 GluGluIleArgAsnIleCysSerAsnSerArgGluLysIleTyAsnArgSerLeuGly 278
QY 802 TCTACTTGTCTCATATGCGCGTACAGAGACTATTGTATACCAAACTGCAGAAACCCA 861
Db 279 SerThrCysHisGlnCysArgGlnLysThrThrAspThrLysThrAsnCysArgAsnPro 298
QY 862 GACTGCTGGGCGTTCGAGGCGAGTTCGTGGCCCTCGCTTCGAAACCGTTATGCTGAA 921
Db 299 AspCysTrpGlyIleArgGlyGlnPheCysGlyProCysLeuArgAsnArgTyrglyglu 318
QY 922 GAGGTCAGGATGCTCTGTCGATCCGAACTGGCATTCGCCCTTCGAGGAAATCTGC 981
Db 319 GluValLysAspAlaLeuAspProAsnTrpHisCysProCysArgGlyIleCys 338
QY 982 AACTGTCAGTTCCTGCGGCGAGAGTGCAGTGGCGTCTGCTGTGTATTATTA 1041
Db 339 AsnCysSerPheCysArgGlnArgAspGlyArgCysAlaThrGlyValLeuValTyLeu 358
QY 1042 GCCAAATATCATGGCTTTGGGAATGTGATGCTACTTGAAGCTGAAACGAGAAATTT 1101
Db 359 AlaLysTyHisGlyPheGlyAsnValHisAlaTyLeuLysSerLeuLysGlnGluPhe 378
QY 1102 GAAATGCAAGCA 1113
Db 379 GluMetGlnAla 382

RESULT 7

Q96BV8 PRELIMINARY; PRT; 220 AA.
ID Q96BV8;
AC Q96BV8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC015124; AAH15124.1; -;
DR InterPro; IPR008972; Cupredoxin.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 220 AA; 25570 MW; C0760EB9AE3899CB CRC64;

Alignment Scores:

Pred. No.: 4,07e-88 Length: 220
Score: 1155.00 Matches: 212
Percent Similarity: 96.80% Conservative: 0
Best Local Similarity: 96.80% Mismatches: 7
Query Match: 58.87% Indels: 0
DB: 2 Gaps: 0

US-10-046-935-2234 (1-1116) x Q96BV8 (1-220)
QY 457 GACTCAGATCAAGAGAGACCGGAGTTCATTCCGGGTGTGTTCCAGAGAGAAC 516
Db 2 AspSerGlnSerArgArgProArgArgThrPheProGlyValAlaSerArgArgAsn 21
QY 517 CCTGAACGAGAGAGTCTGCTCTTACCAGGTCAAGGTCCCGATCCTCGGTCCCTTCAAC 576
Db 22 ProGluArgArgAlaArgProLeuThrArgSerArgSerArgIleLeuGlySerLeuAsp 41
QY 577 GCTTACCCATGCGAGAGCGGTTCGARKMTWGTACATGTTGGTTCAGAGAGAGAG 636
Db 42 AlaLeuProMetGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 61
QY 637 ACCGTGGATGCTACATGAATGAAGATGACCTGCCAGAACCGTTCGATCATCTCC 696
Db 62 ThrValAspGlyTyMetAsnGluAspAspLeuProArgSerArgArgSerArgSer 81
QY 697 GTGACCCCTTCGCATATATAATTCGCCAGTGGAGAAATATACAGAGAGAGTGGAGAAC 756
Db 82 ValThrLeuProHisIleArgProValGluGluIleThrGluGluGluLeuGluAsn 101
QY 757 GTCTGCAGCAATTCGAGAGAGATATATAACCGTTCACCTGGCTCTACTGTCTCA 816
Db 102 ValCysSerAsnSerArgGluLysIleTyAsnArgSerLeuGlySerThrCysHisGln 121
QY 817 TGCCCTCAGAGAGACTATTGATACCAAACTGCAGAAACCCAGACTGCTGGGCGGT 876
Db 122 CysArgGlnLysThrIleAspThrLysThrAsnCysArgAsnProAspCysTrpGlyVal 141
QY 877 CGAGCCAGTTCGTGGCCCTGCTTCGAAACCGTTCGAGAGAGGTTCAGGATGCT 936
Db 142 ArgGlyGlnPheCysGlyProCysLeuArgAsnArgTyrglygluValArgAspAla 161
QY 937 CTGCTGGATCGAACTGCGCATTCGCCCTTCGAGGAACTGCAACTGAGTTCTGTC 996
Db 162 LeuLeuAspProAsnTrpHisCysProCysArgGlyIleCysAsnCysSerPheCys 181
QY 997 CGGACGAGAGATGACCGGTGCGCACTGGGTCTTGTGTATTTAGCCAAATATCATGCG 1056
Db 182 ArgGlnArgAspGlyArgCysAlaThrGlyValLeuValTyLeuAlaLysTyHisGly 201
QY 1057 TTTGGGAATGTCATGCTCTACTTGAAGCCCTGAAACAGGAATTTGAATGCAAGCA 1113
Db 202 PheGlyAsnValHisAlaTyLeuLysSerLeuLysGlnGluPheGluMetGlnAla 220

RESULT 8

Q8NBY9 PRELIMINARY; PRT; 207 AA.
ID Q8NBY9;
AC Q8NBY9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ90653.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahara K.,
RA Masuho Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK075134; BAC11425.1; -;
SQ SEQUENCE 207 AA; 23340 MW; 77AF9BB96EFD0F0F CRC64;

Alignment Scores:

Pred. No.: 3,92e-57 Length: 207
Score: 786.00 Matches: 155

Percent Similarity: 99.36% Conservative: 0
 Best Local Similarity: 99.36% Mismatches: 1
 Query Match: 40.06% Indels: 0
 DB: 2 Gaps: 0

US-10-046-935-2234 (1-1116) x Q8NB9 (1-207)

QY 1 ATGAGCTCGCGCTGCGCGAGAAAGATCTCAGAGTAAAGAACTTAAAGAAATTC 60
 Db 1 MetAspAlaArgValProGlnLysAspLeuArgValLysLysLysLysLysPhe 20
 QY 61 AGATATGTGAAGTTGATTTCCATGGAAACCTCGTCATCTCTGATGACAGTTGACAGC 120
 Db 21 ArgTyrValLysLeuIleSerMetGluThrSerSerSerSerSerSerSerSerSer 40
 QY 121 TTTCCTCTCATATTTTGAACACAGAGCTCGAGTCAGTTCGGGAGGCTGTAGGACC 180
 Db 41 PheAlaSerAspAsnLysPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 60
 QY 181 CGCAGCAGTCGAGCAGCTCTGGACCTCTCAGGCTGCGGATGAAGTTTCCAGCGCGGAGT 240
 Db 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80
 QY 241 ACCAGGGGACCAACCAACAAAGAGAGTCCCGCAGCCCTCAGAGAATTCGTGACT 300
 Db 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100
 QY 301 GATTCCAACCTCCGATTGAGAGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 360
 Db 101 AspSerAsnSerAspSerGluAspGluSerGlyGlyMetAsnPheLeuGluLysArgAlaLeu 120
 QY 361 AATATAAGCAAAACAAAGCAATGCTTGCAGAACTCATGTCTCAATTAGAAGCTTCCCT 420
 Db 121 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140
 QY 421 GGCTCTGCTCGTGAAGACATCCCTCCAGGCTCCGACTCAGCAATCA 468
 Db 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerValSer 156

RESULT 9
 Q96C70

ID Q96C70 PRERIMINARY; PRT; 442 AA.
 AC Q96C70;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE RAM2 protein (Fragment).
 GN Name:RAM2;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1] _SEQUENCE FROM N.A.
 RP SEQUENCE FROM N.A.
 RC TISSUE=Eye;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Villalón D.K., Murzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalek U., Smallos D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human
 RL and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Eye;
 RA Strausberg R.;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC014630; AAH14630.2; -;
 FT NON_TER 1
 SQ SEQUENCE 442 AA; 50745 MW; C57A871694B2C4FD CRC64;

Alignment Scores:

Pred. No.: 3 4e-54 Length: 442
 Score: 752.00 Matches: 172
 Percent Similarity: 57.36% Conservative: 50
 Best Local Similarity: 44.44% Mismatches: 111
 Query Match: 38.33% Indels: 54
 DB: 2 Gaps: 10

US-10-046-935-2234 (1-1116) x Q96C70 (1-442)

QY 22 CAGAAATCTCAGAGTAAAGAACTTAAAGAAATTCAGATATGTGAAGTTGATTC 81
 Db 77 GlnSerAspLeuAsnGlyLysThrAsnProGlu-----ValMet 89
 QY 82 ATGGAACCTCGTCATCTCTGATGACAGTTGTGACAGCTTCTCTGATATTTTGA 141
 Db 90 ValValGluSerAspLeuSerAspGlyLysAlaSerLeuValSerGluGluGlu 109
 QY 142 AACACGAGGCTGCGAGTCAGTTCGGGAAGGCTGAGGACCGCAGCCAGTCGAGCAGTCT 201
 Db 110 AspGluGlu---GluAspLysAlaThrProArgArgSerArgSerArg---ArgSerSer 127
 QY 202 GGACCTCTCAGGTGCGGATGAAGTTTCCAGCCGCGAGTACCAGGGGAGCAACCAACAA 261
 Db 128 IleGlyLeuArgValAlaPheGlnPheProThrLysLysLeuAlaAsnLysProAspLys 147
 QY 262 AAAGCAGACTCCGCCAG----- 279
 Db 148 AsnSerSerSerGluGlnLeuPheSerSerAlaArgLeuGlnAsnGluLysLysThrIle 167
 QY 280 -----CCCTCAGAGAAATCTGTGACTGATTC 306
 Db 168 LeuGluArgLysLysAspCysArgGlnValIleGlnArgGluAspSerThrSerGluSer 187
 QY 307 AACTCCGATTGAGAGATCAAAAGT-----GGAATGAATTTTGGAGAAAGGGCTTTA 360
 Db 188 GluAspAspSerArgAspGluSerGlnGluSerSerAspAlaLeuLysLysArgThrMet 207
 QY 361 AATATAAGCAAAACAAAGCAATGCTTGCAGAACTCATGTCTCAATTAGAAGCTTCCCT 420
 Db 208 AsnIleLysGluAsnLysAlaMetLeuAlaGlnLeuAlaGluLeuAsnSerMetPro 227
 QY 421 GGCTCTGCTCGTGAAGACATCCCTCCAGGCTCCGACTCACAATCAAGGAGACCGCGA 480
 Db 228 AspPhePheProValArg-----ThrProThrSerAlaSerArgLysLysThrValArg 245
 QY 481 AGCGGTACATTCGCGGTGTCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 537
 Db 246 ArgAlaPheSerGluGlyGlnIleThrArgArgMetAsnProThrArgSerAlaArgPro 265
 QY 538 CTTACAGGTCAAGTCCCGGATCTCGGCTCCCTTGAGCTCTACCCATGGRGAGAGCC 597
 Db 266 ProGluLysPhe-----AlaLeuGluAsnPheThrValSerAlaAla 279
 QY 598 AGTTTGTGAKMTWGTACATGTTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 657
 Db 280 LysPheAlaGluLysPheTyrSerPheArgArgArgLysLysLysLysLysLysLys 299
 QY 658 GAAGATGACCTCCAG 717
 Db 300 Glu-----TyrArgArgHisArgIleSer-----SerPhe 310

Query Match:	38..33%	Indels:	54
DB:	2	Gaps:	10
US-10-046-935-2234 (1-1116) x AAH25242 (1-453)			
Qy	22	CAGAAAGATCTCAGAGTAAGAAGAACTTAAAGAAATTCAGATATGTGAAGTTGATTTCC	81
Db	88	GlnSerAspLeuAsnGlyThrAsnProGlu	100
Qy	82	ATGGAACCTCGTCATCCTCTGATGACAGTTGTGACAGCTTTCCTCTGTATAAATTTGCA	141
	:::		
Db	101	ValValGluSerAspLeuSerAspAspGlyLeuAlaSerLeuValSerGluGluGlu	120
	:::		
Qy	142	ACACGAGGCTGCAGTCAGTTCGGGAAGGCTGTAGACCCGACCGCAGTCAGCGACTCT	201
	:::		
Db	121	AspGluGlu---GluAspLysAlaThrProArgSerArgSerArg---ArgSerSer	138
	:::		
Qy	202	GGACCTCTCAGGTGGCGATGAAGTTTCCAGCGCGAGTACACGGGAGCAACCAACAA	261
	:::		
Db	139	IleGlyLeuArgValAlaPheGlnPheProThrLysLysLeuAlaAsnLysProAspLys	158
	:::		
Qy	262	AAAGCAGAGTCCGCCAG	279
	:::		
Db	159	AsnSerSerSerGluGlnLeuPheSerSerAlaArgLeuGlnAsnGluLysLysThrIle	178
	:::		
Qy	280	-----CCCTCAGAGAATTCGTGACTGATTC	306
	:::		
Db	179	LeuGluArgLysLysAspCysArgGlnValIleGlnArgGluAspSerThrSerGluSer	198
	:::		
Qy	307	AACTCCGATTCAAGAAGTAAAGT-----GGAATGAATTTTTGGAGAAAGGGCTTTA	360
	:::		
Db	199	GluAspAspSerArgAspGluSerGlnGluSerSerAspAlaLeuLysArgThrMet	218
	:::		
Qy	361	AATATAAGCAAAACAAAGCAATGCTTGCAAAACCTCATGTCTGAATAGAAAGCTTCCCT	420
	:::		
Db	219	AsnIleLysGluAsnLysAlaMetLeuAlaGlnLeuLeuAlaGluLeuAsnSerMetPro	238
	:::		
Qy	421	GGTCGTTCCGTGGAGACATCCCTCCAGGCTCCGACTCACAATCAAGGACACCGCA	480
	:::		
Db	239	AspPheProValArg-----ThrProThrSerAlaSerArgLysThrValArg	256
	:::		
Qy	481	AGCGGTACATTCGCGGTGTGTCTCCAGGAGA--AACCTCAACCGAGAGCTCGTCC	537
	:::		
Db	257	ArgAlaPheSerGluGlyGlnIleThrArgArgMetAsnProThrArgSerAlaArgPro	276
	:::		
Qy	538	CTTACCAGGTCAAGTCCCGGATCCTCGGTCCCTTGACGCTCTACCCATGGRGAGACC	597
	:::		
Db	277	ProGluLysPhe-----AlaLeuGluAsnPheThrValSerAlaAla	290
	:::		
Qy	598	AGTTTGARKMTWGTACATGTTGGTCAGAAAGGAAGACCTGGATGGCTACATGAAT	657
	:::		
Db	291	LysPheAlaGluPheTySerPheArgArgArgLysThrIleGlyGlyLysCysArg	310
	:::		
Qy	658	GAAGATGACCTGCCAGAACCCCTCGCTCCAGATCATCCGTGACCCCTCCGCATATAAAT	717
	:::		
Db	311	Glu-----TyrArgArgArgHisArgIleSer-----SerPhe	321
	:::		
Qy	718	CGCCGAGTGAAGAAATTACAGAGGAGGTGGAGAACGCTGTGCAGCAATTCGAGAG	777
	:::		
Db	322	ArgProValGluAspIleThrGluGluAspLeuGluAsnValAlaIleThrValArgAsp	341
	:::		
Qy	778	AAGATATATAACCGTTCACTGGGCTCTACTTGTTCATCAATCCGTCAGAGACTATTGAT	837
	:::		
Db	342	LysIleTyAspLysValLeuGlyAsnThrCysHisGlnCysArgGlnLysThrIleAsp	361
	:::		
Qy	838	ACCAAAACAACTGCAGAAACCCAGACTGCTGGGGGCTTCGAGGCCAGTTTCTGTGCCCC	897
	:::		
Db	362	ThrLysThrValCysArgAsnGlnGlyCysGlyValArgGlyGlnPheCysGlyPro	381
	:::		
Qy	898	TGCTTCGAAACCGTTATGGTGAAGAGGTGAGGATGCTCTGCTGGATCCGAACATGGCAT	957
	:::		
Db	382	CysLeuArgAsnArgTyrGlyGluAspValArgSerAlaLeuLeuAspProAspTspVal	401
	:::		

QY	958	TGCCCGCCCTGTGAGGAATCTGCAACTGCAGTTTCTGCCGCGCAGCGAGATGACCGTGT	101
Db	402	CysProProcysargGlyIleCysAsnCysSerTyrCysargIysargAepGlyArgCys	421
QY	1018	GCACACTGGGCTCTGTGTATTTAGCCAAATATCATGCTTTCGGAAATGTCATGCTCTAC	1077
Db	422	AlaThrGlyIleLeuIleHisLeuAlaIasPheTyrGlyTyrAspAsnValLysGluTyr	441
QY	1078	TTGAAAAGCTGAAACACGAA 1098	
Db	442	LeuGluSerLeuGlnIysGlu 448	

RESULT 12			
Q96GN5	PRELIMINARY;	PRT;	454 AA.
ID	Q96GN5	Q81XN5;	
AC	01-DRC-2001 (TrEMBLrel. 19, Created)		
DT	01-MAR-2004 (TrEMBLrel. 26, Last sequence update)		
DT	01-OCT-2004 (TrEMBLrel. 28, Last annotation update)		
DE	Transcription factor RAM2 splice variant c.		
GN	Name=RAM2;		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
NCBI_TaxID=9606;			
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Lung, and Muscle;		
RK	MEDLINE=22388257; PubMed=12477932;		
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,		
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavani T.L., Scheetz T.E.,		
RA	Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,		
RA	Bosak S.A., McSwain P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards D., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalon S.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko I., Bouffard G.G.,		
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,		
RA	Krzywinski M.I., Skalska U., Smullus D.E., Schnerch A., Schein J.E.,		
RA	Jones S.J., Marra M.A.;		
RT	"generation and initial analysis of more than 15,000 full-length human		
RT	and mouse cDNA sequences.";		
RT	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Muscle;		
RC	Strausberg R.;		
RA	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RA	Cathomen T., Wang K.H., Oswald W.B., Weitzman M.D.;		
RL	Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Lung;		
RA	Strausberg R.;		
RL	Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC009352; AAH09352.2; -		
DR	EMBL; AV161169; AAO1751.1; -		
DR	EMBL; BC039823; AAH39823.1; -		
DR	SEQUENCE 454 AA; 52206 MW; 1FAFEC7DBD763C85 CRC64;		

Alignment Scores:			
Pred. No.:	3,42e-54	Length:	454
Score:	752.00	Matches:	172
Percent Similarity:	75.36%	Conservative:	50
Best Local Similarity:	44.44%	Mismatches:	111


```

RESULT 14
RAH32576
ID AAH32576 PRELIMINARY; PRT; 453 AA.
AC AAH32576;
DT 29-MAR-2004 (TrEMBLrel. 27, Created)
DT 29-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 29-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Transcription factor RAM2.
GN RAM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Rubin A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Rosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.W., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC032576; AAH32576.1; -.
SQ SEQUENCE 453 AA; 52018 MW; 6AA046560F31DABB CRC64;

Alignment Scores:
Pred. No.: 1,95e-53 Length: 453
Score: 743.00 Matches: 171
Percent Similarity: 57.11% Conservative: 50
Best Local Similarity: 44.19% Mismatches: 112
Query Match: 37.87% Indels: 54
DB: 2 Gaps: 10

US-10-046-935-2234 (1-1116) x AAH32576 (1-453)
QY 22 CAGAAAGATCTCAGAGTAAAGAAAGAACTTAAAGAAATTCAGATATGTGAAGTTGATTTC 81
Db 88 GlnSerAspLeuAsnGlyLysThrAsnProGlu-----ValMet 100
QY 82 ATGGAACCTCGTCATCTCTGATGACAGTTGTGACAGTTGTGCTTCTGATATTTTGA 141
Db 101 ValValGluSerAspLeuSerAspGlyLysAlaSerLeuValSerGluGluGlu 120
QY 142 AACACGAGGCTCAGTTCAGTTCGGGAAGCTGTAGGACCGGACGAGTGCAGGACTCT 201
Db 121 AspGluGlu---GluAspLysAlaThrProArgArgSerArg---ArgSer 138
QY 202 GCACCTCTCAGGTCGCGATGAGTTTCCAGGCGGAGTACCAGGGGAGCAACCAACAA 261
Db 139 IleGlyLeuArgValAlaPheGlnPheProThrLysLeuAlaAsnLysProAspLys 158
QY 262 AAAGCAGAGTCCCGCAG----- 279

```

```

Db 159 AsnSerSerSerGluGlnLeuPheSerSerAlaArgLeuGlnAsnGluLysLysThrIle 178
QY 280 -----CCCTCAGAGAAATTCGTGACTGATCC 306
Db 179 LeuGluArgLysLysAspCysArgGlnValIleGlnArgGluAspSerThrSerGluSer 198
QY 307 AACTCCGATTTCAGAAAGT-----GGAATGAATTTTTCGAGAAAGGCTTTA 360
Db 199 GluAspAspSerArgAspGluSerGlnGluSerSerAspAlaLeuLysArgThrMet 218
QY 361 AATATAAGCAAAACAAAGCAATGTTCCAAACTCATGCTGAATTAAGAAAGCTCCCT 420
Db 219 AsnIleLysGluAsnLysAlaMetLeuAlaGlnLeuLeuAlaGluLeuAsnSerMetPro 238
QY 421 GGCTCGTTCCGTGGAGACATCCCTCCAGGCTCCGACTCCACATCACATCAGGAGACCGGA 480
Db 239 AspPhePheProValArg-----ThrProThrSerAlaSerArgLysLysThrValArg 256
QY 481 AGGCGTACATTCCTCCGGTGTTCCTCCAGGAGA---AACCTCGAAGCGAGAGCTCGTCT 537
Db 257 ArgAlaPheSerGluGlyGlnIleThrArgArgMetAsnProThrArgSerAlaArgPro 276
QY 538 CTTACAGGTCAGGTCCTCCGATCCTCGGTCCCTTGGAGCTCTACCCATGGRGAGGCC 597
Db 277 ProGluLysPhe-----AlaLeuGluAsnPheThrValSerAlaAla 290
QY 598 AGTTTGGARKMTWGTACATGTTGTTGAGAGAGAGACCGTGGTGGCTACATGAAT 657
Db 291 LysPheAlaGluGluPheThrPheArgArgLysThrIleGlyLysCysArg 310
QY 658 GAAGATGACCTGCCAGAGAGCGTCCGATCATCCGATCATCCGATCCGATCCGATTAAT 717
Db 311 Glu-----TyrArgArgHisArgIleSer-----SerPhe 321
QY 718 GCGCCAGTGGAGAAATTCAGAGAGAGAGTGGAGAGAGCTGCGAGCAATTCCTCAGAG 777
Db 322 ArgProValGluAspIleThrGluGluAspLeuGluAsnValAlaIleThrValArgAsp 341
QY 778 AAGATATATACCGTTCACTGGGCTCTACTTCTCATCAATGCGCTCAGAGACTATTGAT 837
Db 342 LysIleCysAspLysValLeuGlyAsnThrCysHisGlnCysArgGlnLysThrIleAsp 361
QY 838 ACCAAAACAAATCGCAGAAACCCAGACTCTGGGGCGTTCGAGGCGCAGTCTGTGSCCC 897
Db 362 ThrLysThrValCysArgAsnGlnGlyCysCysGlyValArgGlyGlnPheCysGlyPro 381
QY 898 TGCCTTCGAACCGTTTGGTGAAGAGGTGAGGATGCTCTGCTGGATCCGAACTGGCAT 957
Db 382 CysLeuArgAsnArgTyrGlyGluAspValArgSerAlaLeuLeuAspProAspTrpVal 401
QY 958 TGCCCGCTTGTGAGAGAACTGCAACTGCAGTTTCTGCCGCGCAGCAGATGGACGGTCT 1017
Db 402 CysProProCysArgGlyLysCysAsnCysSerTyrCysArgLysArgAspGlyArgCys 421
QY 1018 GCGACTGGGGTCTTGTGTATTAGCCAAATATCATGGCTTTGGGAATGTGATCGCTAC 1077
Db 422 AlaThrGlyIleLeuIleHisLeuAlaLysPheTyrGlyTyrAspAsnValLysGluTyr 441
QY 1078 TTGAAAAGCTGAAACAGAA 1098
Db 442 LeuGluSerLeuGlnLysGlu 448
RESULT 15
Q922M5 PRELIMINARY; PRT; 438 AA.
AC Q922M5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CDNA sequence BC006933.
GN Name=BC006933;
OS Mus musculus (Mouse).

```

Qy	442	CCCTCCAGGCTCCGACTCAAACTACAGAGACCGCAAGCGTCAATCTCCCGGTGTT	501
Db	234	---ThrProSerAlaSerArgArgThrProArgArgAlaPheSerGluGlyGln	252
Qy	502	GCTTCAGGAGA---AACCCCTGAACGAGAGCTCGTCTCT	537
Db	253	IleThrArgArgMetAsnProThrArgSerAlaArgProGluLysPheAlaLeuGlu	272
Qy	538	---CTTACCAGGTCAAAGTCCCGATNCTCGGGTCCCTTGACGCTTCACCCATGGRGAGA	594
Db	273	AsnPheThrPheSerAlaThrLysLeu-----	281
Qy	595	GCACGTTTKARWMTWGTACATGTTGGTCAGAAAGAGAGACCGTGGATGGCTACATG	654
Db	282	-----ThrGluGluLeuTyrSerPhe---ArgArgArgLysThrIleSerGly-----	296
Qy	655	AATGAAGATGACCTGCCAGAACCGTCGCTCCAGATCATCCGTGACCCCTCCGCATATA	714
Db	297	-----GlyLysCysGlnThrTyrArgArgHisArgIleSer-----Ser	309
Qy	715	ATTGCCCCGTGGAAGAAATACAGAGAGAGAGTGGAGAACGCTTCGAGCAATCTCGA	774
Db	310	PheArgSerValLysAspIleThrGluGluAspLeuGluAsnIleAlaIleThrValArg	329
Qy	775	GAGAAGATATAAACCGTTCACCTGGGCTCTACTGTGCATCAATCAATGCCCTCAGAAGACTATT	834
Db	330	AspLysValTyrAspLysValLeuGlyAsnThrCysHisGlnCysArgGlnLysThrIle	349
Qy	835	GATACCAAAACAACTCAGAAACCCAGACTGCTGGGGCGCTTCAGAGCCAGTTCGTGGC	894
Db	350	AspThrLysThrValCysArgAsnGlnSerCysGlyGlyValArgGlyGlnPheCysGly	369
Qy	895	CCCTGCTTCGAAACCCCTTATGTGTGAAGAGGTCAAGGATGCTCTGTGGATCCGAATCG	954
Db	370	ProCysLeuArgAsnArgTyrGlyGluAspValArgThrAlaLeuLeuAspProLysTrp	389
Qy	955	CATTGCCGCTTGTCCAGGAACTCGAACTGCAGTTCCTCCGGCAGCAGATGGACGG	1014
Db	390	ThrCysProProCysArgGlyIleCysAsnCysSerTyrCysArgArgAspGlyArg	409
Qy	1015	TGTCGCACTGGGGTCTTGTTGATTATTAGCCAAATATCATGCTTGGGAATGTCATGCC	1074
Db	410	CysAlaThrGlyIleLeuIleHisLeuAlaLysPheTyrGlyTyrAspAsnValLysGlu	429
Qy	1075	TACTTGAAGAGCTGAAACAGGAA	1098
Db	430	TyrLeuGluSerLeuGlnLysGln	437

Search completed: October 12, 2004, 15:18:39
Job time : 184.5 secs

```

Qy 955 CATTCGCCGCTTGTCGAGGAATCTGCAACTGCAGTTTCTGCCGCGACCGAGATGGACGG 1014
Db 390 ThrCysProCysArgGlyIleCysAsnCysSerTyrCysArgArgAspGlyArg 409
Qy 1015 TGTGCGACTGGGTCCTTGTTGTTATTAGCCAAATATCATGGCTTTGGAAATGTCATGCC 1074
Db 410 CysAlaThrGlyIleLeuIleHisLeuAlaLysPheTyrGlyTyrAspAsnValysGlu 429
Qy 1075 TACTTGAAAAGCCTGAAACAGGAA 1098
Db 430 TyrLeuGluSerLeuGlnLysGln 437

Search completed: October 12, 2004, 15:18:39
Job time : 184.5 secs

```

Blank

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 12, 2004, 15:04:25 ; Search time 112 Seconds
(without alignments)
7148.962 Million cell updates/sec

Title: US-10-046-935-2234

Perfect score: 1962

Sequence: 1 atggacgtcgcgcgtgtgc.....aattgaaatgaagcataa 1116

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 4004546

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-O=/cpn2_1/USPTO spool_p/US10046935/runat 12102004 155225 15233/app_query.fasta_1.1287
-DB=A_Geneseq 23Sep04 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITIS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10046935 @CGN 1.1 1.77 @runat 12102004 155225 15233 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq 23Sep04:.*
1: Geneseq1980s:.*
2: Geneseq1990s:.*
3: Geneseq2000s:.*
4: Geneseq2001s:.*
5: Geneseq2002s:.*
6: Geneseq2003as:.*
7: Geneseq2003bs:.*
8: Geneseq2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1943	99.0	371	5	Abb05376 Human col
2	1929	98.3	371	4	Aab94473 Human pro
3	1928	98.3	371	6	Abu56624 Lung can
4	1929	98.3	371	7	Adf76612 Novel hum
5	1929	98.3	371	7	Adn38808 Novel hum
6	1929	98.3	371	8	Adn38808 Cancer/an
7	1929	98.3	371	8	Adg09178 Human CDC
8	1929	98.3	407	3	Aab43720 Human can
9	1924	98.1	397	4	Aag74880 Human col
10	1924	98.1	397	6	Ada98284 Human sec
					Ada44119 Human sec

11	1924	98.1	397	7	ADC20449	Adc20449 Human sec
12	1924	98.1	397	7	ADF10748	Adf10748 Human sec
13	1879.5	95.8	450	4	AAB93122	Aab93122 Human pro
14	1879.5	95.8	450	8	ADL91540	Adl91540 Human imm
15	1879.5	95.8	450	8	ADN06070	Adn06070 Antipsori
16	1879.5	95.8	450	8	ADO20297	Ado20297 Human PRO
17	1879.5	95.8	450	8	ADQ09208	Adq09208 Human CDC
18	1879.5	95.8	450	8	ADQ09176	Adq09176 Human CDC
19	1656	84.4	347	5	AAG79357	Aag79357 CGA7 pref
20	1656	84.4	347	5	ADN39876	Adn39876 Cancer/an
21	1555	79.3	298	8	ADN99718	Adn99718 Novel hum
22	1520.5	77.5	320	3	AAB23164	Aab23164 Human col
23	1520.5	77.5	320	5	AAG79356	Aag79356 CGA7. 8/2
24	1520.5	77.5	320	7	AD848023	Ad848023 Human CGA
25	852	43.4	174	6	ABU71082	Abu71082 Human adi
26	786	40.1	207	4	AAM93560	Aam93560 Human pol
27	786	40.1	207	8	ADL31298	Adl31298 Human pro
28	752	38.3	420	7	ADM04753	Adm04753 Human pro
29	563.5	28.7	110	4	ABG22773	Abg22773 Novel hum
30	535	27.3	89	3	AAB32062	Aab32062 Human sec
31	498.5	25.4	222	4	ABG03573	Abg03573 Novel hum
32	424.5	21.4	344	4	ABG03571	Abg03571 Novel hum
33	346	17.6	71	4	ABG22772	Abg22772 Novel hum
34	241.5	12.3	102	3	AAB32061	Aab32061 Human sec
35	195	9.9	40	3	AAB32010	Aab32010 Human sec
36	171.5	8.7	268	4	AAB94363	Aab94363 Human pro
37	146	7.4	875	8	ADD062965	Add062965 Transcrip
38	138	7.0	786	7	ADD30845	Add30845 Plant yie
39	138	7.0	786	8	ADI43825	Adi43825 Plant tra
40	132.5	6.8	570	8	ADN72417	Adn72417 Thale cre
41	131.5	6.7	906	8	ADO61723	Ado61723 Transcrip
42	129	6.6	754	5	AAB24596	Aab24596 Human SR-
43	129	6.6	754	7	ADG10664	Adg10664 Human STA
44	129	6.6	754	8	ADH09504	Adh09504 Human hos
45	127	6.5	448	7	ADG10780	Adg10780 Human STA

ALIGNMENTS

RESULT 1

ABB05376
ID ABB05376 standard; protein; 371 AA.

XX AC ABB05376;

XX DT 08-APR-2002 (first entry)

XX DE Human colon tumour antigen polypeptide SEQ ID NO:2235.

XX DE Human; colon cancer; colon tumour antigen; cytostatic; vaccine;

XX KW Human; colon tumour metastatic antigen; diagnosis.

XX OS Homo sapiens.

XX PN WO200196388-A2.

XX PD 20-DEC-2001.

XX PF 08-JUN-2001; 2001WO-US018557.

XX PR 09-JUN-2000; 2000US-0210899P.

XX PR 20-FEB-2001; 2001US-0270216P.

XX PA (CORI-) CORIXA CORP.

XX PI Jiang Y, Harlocker SL, Secrist H;

XX PI WPI; 2002-114514/15.

XX PT Novel isolated colon tumor polynucleotide differentially expressed in colon tumor or colon metastatic tumor and polypeptides encoded by them, useful for inhibiting development of cancer in patient.

XX

PS Claim 2; SEQ ID NO 2235; 105pp; English.

XX ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)
 CC which were isolated from human colon tumour and colon metastatic tumour
 CC cDNA libraries. (I) have cytosolic activity and can be used in vaccine
 CC production. (I) can be used for stimulating and/or expanding T cells
 CC specific for a tumour protein on contact with the T cells. They are also
 CC useful for inhibiting the development of cancer in a patient. (I) can be
 CC used as probes or primers for nucleic acid hybridisation, for preparing
 CC mutant species primers, or primers for use in genetic constructions. (I)
 CC can be used in the diagnosis of a colon tumour. The present sequence
 CC represents a human colon tumour antigen amino acid sequence which is
 CC specifically claimed in the present invention

XX Sequence 371 AA;

Alignment Scores:

Pred. No.: 6,078-196 Length: 371
 Score: 1943.00 Matches: 371
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.03% Indels: 0
 DB: 5 Gaps: 0

US-10-046-935-2234 (1-1116) x ABB05376 (1-371)

QY 1 ATGGAGCTCGCGGTGCGGAGAAAGATCTCAGAGTAAGAAGAACTTAAGAATTC 60
 Db 1 MetAspAlaArgValProGlnLysAspLeuArgValLysLysAsnLeuLysPhe 20
 QY 61 AGATATGTGAAGTTGATTTCCATGGAAACCTCGTCATCTCTGATGACAGTTGTGACAC 120
 Db 21 ArgTyrValLysLeuLeuSerMetGluThrSerSerSerSerAspSerCysAspSer 40
 QY 121 TTGTCTCTGATAATTTTCCAAACACAGAGCTGCAGTTCAGTTCGGGAAGCTGTAGGAC 180
 Db 41 PheAlaSerAspAsnPheAlaSerThrArgLeuGlnSerValArgGluGlyCysArgThr 60
 QY 181 CGCAGCAGTGCAGGCACTCTGGACCTCTCAGGTGGCGATGAAGTTTCCAGCGCGAGT 240
 Db 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80
 QY 241 ACCAGGGAGCAACCAACAAAAGCAGAGTCCCGCCAGCCCTCAGAGAATTTGTGACT 300
 Db 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100
 QY 301 GATTCACCTCCGATTCAGAGATGAAGTGAAGTAATTTTTCGAGAAAGGGCTTTA 360
 Db 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuLysArgAlaLeu 120
 QY 361 AATATAAGCAAAACAAAGCAATGCTTGCAAAACCTCATGCTGAATAGAAAGCTTCCCT 420
 Db 121 AsnLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140
 QY 421 GGCTCGTTCCGTGGAGACATCCCTCCAGGCTCCGACTCACAATCAAGGAGACCGCGA 480
 Db 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160
 QY 481 AGCGGTACATTCGCGGTGCTGCTCCAGGAGAAACCTTGAACGGAGAGCTGCTCCTTT 540
 Db 161 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgArgAlaArgProLeu 180
 QY 541 ACCAGTCAAGTTCGCGGATCTTCGGGTCCCTTGACGCTCTACCCATGGGAGAGCACT 600
 Db 181 ThrArgSerArgSerArgLysLeuGlySerLeuAspAlaLeuProMet***ArgAlaSer 200
 QY 601 TTKGARKMTWTFACATGTTGGTGAGAAAGAGAGAACCGTGGATGGCTACATGAATGA 660
 Db 201 ***Glu*****TyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220
 QY 661 GATGACCTGCCAGAGACCGTCCCTCCAGATCATCCGTGACCTTCCGGCATATAATTCG 720
 Db 221 AspAspPheProArgSerArgArgSerArgSerValThrLeuProHisLysLysLeuArg 240

QY 721 CCAGTGGAGAAATTTACAGAGGAGGAGTTGGAGAACTCTGCAGCAATTTCTCAGAGAAG 780
 Db 241 ProValGluGluLysThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLys 260
 QY 781 APATATAACCGTCTACTGGCTCTACTTGTCTATCAATGCGCTCAGAGAATTTGATACC 840
 Db 261 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 280
 QY 841 AAAACAACCTCAGAAACCCAGACTGCTGGGCGCTTCAGAGCCAGTCTGTCGCCCTCG 900
 Db 281 LysThrAsnCysArgAsnProAspCysTrpGlyValArgGlnPheCysGlyProCys 300
 QY 901 CTTGAAACCGTTATGCTGAGAGGTCAGAGGATGCTCTGCTGGATCCGAACTGGCATTC 960
 Db 301 LeuArgAsnArgTyrGlyGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 320
 QY 961 CGCCTTGTGAGGAATCTGAACTGCGAGTTTCTGCCGCGCAGCAGATGAGCGTGTGG 1020
 Db 321 ProProCysArgGlyLeuCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 340
 QY 1021 ACTGGGCTCTGCTGTATTTAGCCAAATATCATGCTTTGGGAATGCTGATGCTACTTG 1080
 Db 341 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 360
 QY 1081 AAAAGCCTGAAACAGCAATTTGAAATGCAAGCA 1113
 Db 361 LysSerLeuLysGlnGluPheGluMetGlnAla 371
 RESULT 2
 AAB94473
 ID AAB94473 standard; protein; 371 AA.
 XX AAB94473;
 XX 26-JUN-2001 (first entry)
 DE Human protein sequence SEQ ID NO:15139.
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX Homo sapiens.
 XX EP1074617-A2.
 PD 07-FEB-2001.
 XX 28-JUL-2000; 2000EP-00116126.
 PF 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX (HELI-) HELIX RES INST.
 XX Ota T, Isogai T, Nishikawa T, Hayaishi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.
 XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
 XX Claim 8; SEQ ID NO 15139; 2537pp + Sequence Listing; English.
 PS The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602

PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.

XX PS Claim 27; Page 353; 453pp; English.

XX CC The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABUS6408-ABUS6745 represent lung cancer-associated polypeptides of the
 CC invention

XX SQ Sequence 371 AA;

Alignment Scores:

Pred. No.: 1,84e-194 Length: 371
 Score: 1929.00 Matches: 364
 Percent Similarity: 98.11% Conservative: 0
 Best Local Similarity: 98.11% Mismatches: 7
 Query Match: 98.32% Indels: 0
 DB: 6 Gaps: 0

US-10-046-935-2234 (1-1116) x ABUS6624 (1-371)

QY 1 ATGGAGCTCGCGCGCGCGGAGAAAGATCTCAGAGTAAAGAAAGAAATTC 60
 DB 1 MetAspAlaArgValProGlnLysAspLeuArgValLysLysAsnLeuLysPhe 20
 QY 61 AGATATGTGAAGTGTATTCATGGAAACCTGTCATCTCGATGACAGTTGTGACAGC 120
 DB 21 ArgTyrValLysLeuLysSerMetGluThrSerSerSerSerAspAspSerCysAspSer 40
 QY 121 TTGTCTCTGATATTTTCAACACAGAGCTGCAGTTCAGTTCGGGAGGCTGTAGACC 180
 DB 41 PheAlaSerAspAsnPheAlaSerThrArgLeuGlnSerValArgGluGlyCysArgThr 60
 QY 181 CGCAGCCAGTGCAGGCACTCTGGACCTCTCAGGTCGCGGATGAAGTTTCCAGCGCGAGT 240
 DB 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80
 QY 241 ACAGGGGAGCAACCAACAAAGAGAGTCCGCGCCAGCTTCAGAGAAATCTGTGACT 300
 DB 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100
 QY 301 GATTCCAACTCCGATTCCAGAGATGAAGTGGAAATTTTGGAGAAAGGGCTTTA 360
 DB 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuLysArgAlaLeu 120
 QY 361 AATATAAGCAAAACAAAGCAATGCTTGCAAAACCTCATGCTGAATTAGAAGCTTCCCT 420
 DB 121 AsnLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140
 QY 421 GAGTCGTTCCGTGGAGACATCCCTCCAGGCTCCGACTCACAATCAAGGACCGCA 480
 DB 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160
 QY 481 AGCGGTACATTCCTCGGTGTGTCTCCAGGAGAAACCTGAAACGAGAGCTCGCTCTCT 540
 DB 161 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgAlaArgProLeu 180

QY 541 ACCAGTCAAGTCCCGATCCTCGGTCCCTTACCGCTTACCCATGGAGAGCCAGT 600
 DB 181 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGlu 200
 QY 601 TTKARKMTMGTACATGTTGTGTGAGAAAGAGAGACCGTGGATGGCTACATGAATGAA 660
 DB 201 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220
 QY 661 GATGACCTCCAGAGCGCTCGCTCCAGATCATCGTGACCTCCGATATTAATTCG 720
 DB 221 AspAspLeuProArgSerArgSerArgSerValThrLeuProHisIleIleArg 240
 QY 721 CCAGTGGAGAAATACAGAGGAGGAGTTCGAGAAACGCTCGCAGCAATTCGAGAGAAG 780
 DB 241 ProValGluGluIleThrGluGluLeuGluAsnValCysSerAsnSerArgGluLys 260
 QY 781 ATATATAACCGTTCATCTGGCTCTACTGTCTCATCAATGCGCTCAGAAGACTATTGATACC 840
 DB 261 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 280
 QY 841 AAAACAACTGCAGAAACCCAGACTGCTGGGCGCTTCGAGGCGCAGTTCTGTGCCCCCTGC 900
 DB 281 LysThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCys 300
 QY 901 CTTGAAACCGTTATGTGTGAAGAGGTTCAGGGATGCTCTGCTGGATCCGAACTGGCATTCG 960
 DB 301 LeuArgAsnArgTyrGlyGluGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 320
 QY 961 CCGCTTGTTCGAGGAATCTGAACTGCAGTTCCTGCGCGCAGCAGATGCGGTGTGG 1020
 DB 321 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 340
 QY 1021 ACTGGGCTCTGTGTATTAGCCAAATATCATGGCTTTGGGAATGTGATGCTACTTGTG 1080
 DB 341 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 360
 QY 1081 AAAAGCCTGAAACAGGAATTTGAAATGCAAGCA 1113
 DB 361 LysSerLeuLysGlnGluPheGluMetGlnAla 371

RESULT 4

ADF76612

ID ADF76612 standard; protein; 371 AA.

XX ADF76612;

XX AC

XX 26-FEB-2004 (first entry)

XX DT

XX DE

XX Novel human secreted and transmembrane protein SeqID.286.

XX human; PRO; membrane bound protein; membrane bound receptor;

XX cell proliferation; cell migration; cell differentiation;

XX mitogenic factor; survival factor; cytotoxic factor; receptor;

XX differentiation factor; neuropeptide; hormone; cell receptor;

XX receptor-ligand interaction; cytostatic; chondrocyte; tumour.

XX OS

XX Homo sapiens.

XX XX

XX WO2003072035-A2.

XX PN

XX 04-SEP-2003.

XX PD

XX 21-FEB-2003; 2003WO-US005241.

XX XX

XX 22-FEB-2002; 2002US-0359461P.

XX PR

XX (GETH) GENENTECH INC.

XX PA

XX Bodary SC, Clark H, Hunte B, Jackman JK, Schoenfeld JR;

XX PI Williams PM, Wood WI, Wu TD;

XX WPI; 2003-721702/68.

XX DR

DR N-PSDB; ADF76611.
XX
PT New PRO polypeptides, useful for diagnosing and treating an immune
PT related disorder, e.g. systemic lupus erythematosus, rheumatoid
PT arthritis, osteoarthritis, juvenile chronic arthritis, thyroiditis or
PT diabetes mellitus.
XX
PS Claim 10; SEQ ID NO 286; 918pp; English.
XX
CC This invention relates to novel nucleic acids encoding human PRO secreted
CC and transmembrane proteins. Extracellular proteins play important roles
CC in the formation, differentiation and maintenance of multicellular
CC organisms. The fate of many individual cells (for example proliferation,
CC migration or differentiation) is typically governed by information
CC received from other cells and the immediate environment. The information
CC is often transmitted by secreted polypeptides (for example mitogenic
CC factors, survival factors, cytotoxic factors, differentiation factors,
CC neuropeptides and hormones) which are received and interpreted by diverse
CC cell receptors or membrane bound proteins. These membrane bound proteins
CC and receptors may be of use as pharmaceutical and diagnostic agents, such
CC as in the blocking of receptor-ligand interactions. The current invention
CC provides the amino acid sequences of novel human membrane bound receptors
CC and proteins, along with the cDNA sequences encoding them. The novel
CC proteins of the invention may have cytostatic activities through the
CC stimulation of chondrocytes. The nucleic acids of the invention may be
CC useful for the manufacture of a medicament for diagnosing or treating a
CC tumour in a mammal. In addition, they may be useful for measuring or
CC detecting the expression of a tumour associated gene. The present
CC sequence is the amino acid sequence of a human PRO protein of the
CC invention.
XX
SQ Sequence 371 AA;
Alignment Scores:
Pred. No.: 1,84e-194 Length: 371
Score: 1929.00 Matches: 364
Percent Similarity: 98.11% Conservative: 0
Best Local Similarity: 98.11% Mismatches: 7
Query Match: 98.32% Indels: 0
DB: 7 Gaps: 0
US-10-046-935-2234 (1-1116) x ADF76612 (1-371)
QY 1 ATCGAGCTCCGCGTCCGCGAGAAAGATCTCAGAGTAAAGAGAACTTAAAGAAATTC 60
Db 1 MetAspAlaArgValProGlnLysAspLeuArgValLysAsnLeuLysLysPhe 20
QY 61 AGATATGTGAAGTTGATTTCCATGGAACCTCGTCATCCTCTGATGACAGTGTGACAGC 120
Db 21 ArgTyrValLysLeuLysSerMetGluThrSerSerSerSerSerSerSerSerSer 40
QY 121 TTTGCTTCTGATAAATTTTGCAACACGAGGCTCGAGTCAGTTCGGGAAGGCTGAGACC 180
Db 41 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 60
QY 181 CGCAGCAGTCGAGGCACTCTGGACCTCTCAGGCTGGCGATGAAGTTTCAGCGCGGAGT 240
Db 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80
QY 241 ACCAGGGGACCAACCAAAAGCAGAGTCCGCGCAGCCCTCAGAGAATTCGTGACT 300
Db 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100
QY 301 GATTCCAACCTCCGATTCAGAGAATGAAGTGGAAATGATTTTGGAGAAAGGGCTTTA 360
Db 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 120
QY 361 AATATAAGCAAAACAAAGCAATGCTTGCAGAACTCATGTCTGAATAGAAAGCTCCCT 420
Db 121 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140
QY 421 GCGTCGTTCCGTGGAGACATCCCTCCAGGCTCCGACTCAAACTCAAGGAGACCGCGA 480

Db 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160
QY 481 AGCGGTACATTCGCGGTGTTGCTTCAGAGAGAAACCTGAACGAGAGCTGCTCTTT 540
Db 161 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgArgAlaArgProLeu 180
QY 541 ACCAGGTCAAAGTCCCGGATCCCTCGGTCTCTCACCTCTTACCCTACCGAGAGCCAGT 600
Db 181 ThrArgSerArgSerArgLleLeuGlySerLeuAspAlaLeuProMetGluGluGlu 200
QY 601 TTKGARKMTWMGTACATGTTGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 201 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220
QY 661 GATGACCTGCCAG 720
Db 221 AspAspLeuProArgSerArgArgSerValThrLeuProHisLleArg 240
QY 721 CCAGTGAAGAAATTCAG 780
Db 241 ProValGluGluLeuThrGluGluGluGluGluGluGluGluGluGluGluGluGlu 260
QY 781 ATATATAACCGTTCACGTGGCTCTACTGTGATCAATGCCGTCAAGAGACTATTGATACC 840
Db 261 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrLleAspThr 280
QY 841 AAAACAACTGCAGAACCCAGAGCTCTGGGGCGTTGAGCCAGGTTCTGTGGCCCTGC 900
Db 281 LysThrAsnCysArgAsnProAspCysTyrGlyValArgGlyGlnPheCysGlyProCys 300
QY 901 CTTTCGAACCGTTCATGTTGAAGAGAGTCTGCTGGATCCGATCCGAACTGGCATTGC 960
Db 301 LeuArgAsnArgTyrGlyGluGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 320
QY 961 CCGCTTGTTCGAGAAATTCGAACTCGAGTTCTCGCGCAGCAGAGATGACCGTGTGCG 1020
Db 321 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 340
QY 1021 ACTGGGGTCTCTGTTATTTAGCCAAATATCATCGCTTGGGAATGTGCATGCTACTTG 1080
Db 341 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 360
QY 1081 AAAAGCTGAAACAGAGAAATTTGAAATGCAAGCA 1113
Db 361 LysSerLeuLysGlnGluPheGluMetGlnAla 371
RESULT 5
ADN38808
ID ADN38808 standard; protein; 371 AA.
XX
AC ADN38808;
XX
XX 17-JUN-2004 (first entry)
DT
DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:126.
XX
XX Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; lechaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnary; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
XX WO2003042661-A2.
XX
XX 22-MAY-2003.
XX
XX 13-NOV-2002; 2002WO-US036810.
PF
XX 13-NOV-2001; 2001US-0350666P.
XX

DE	Human CDCA7/JPO1 protein SEQ ID NO:363.
XX	
KW	thanatos-associated protein; THAP; THAP responsive gene; THAP family;
KW	THAP responsive element; angiogenesis; inflammation; metastasis; cancer;
KW	apoptosis; cardiovascular disease; neurodegenerative disease; chemokine;
KW	antiangiogenic; antiinflammatory; cardiovascular; cytostatic;
KW	neuroprotective; osteopathic; THAP modulator; THAP synthesis modulator;
KW	human.
XX	
OS	Homo sapiens.
XX	
PN	WO2004055050-A2.
XX	
PD	01-JUL-2004.
XX	
PF	10-DEC-2003; 2003WO-IB006434.
XX	
PR	10-DEC-2003; 2002US-0432699P.
PR	03-JUL-2003; 2003US-0485027P.
XX	
PA	(ENDO-) ENDOCYTE SAS.
XX	(CNRS) CNRS CENT NAT RECH SCI.
XX	
PI	Girard J, Amalric F, Rousigne M, Clouaire T;
XX	
DR	WPI; 2004-525034/50.
DR	N-PSDB; ADQ09179.
XX	
XX	Example 45; SEQ ID NO 363; 612pp; English.
XX	
CC	The present invention describes a method for modulating the expression of
CC	a thanatos (death)-associated protein (THAP) responsive gene. The method
CC	comprises modulating the interaction of a THAP-family polypeptide or its
CC	biological fragment with a nucleic acid, and so enhancing or repressing
CC	the expression of the THAP responsive gene. Also described: (1) a method
CC	of modulating the expression of a gene responsive to a THAP/chemokine
CC	complex; (2) a pharmaceutical composition comprising a THAP responsive
CC	element in a pharmaceutical carrier; (3) a transcription factor decoy
CC	consisting essentially of a THAP responsive element; (4) a cell
CC	comprising a transcription factor decoy described above; (5) methods of
CC	modulating the interaction between a nucleic acid and a THAP-family
CC	polypeptide or its biological fragment, or a nucleic acid and a
CC	THAP/chemokine complex; (6) a vector packaging cell line comprising a
CC	cell comprising a viral vector which comprises a promoter operably linked
CC	to a nucleic acid encoding a THAP-family polypeptide or its biological
CC	fragment; (7) a method of constructing a cell which expresses a
CC	recombinant THAP-family polypeptide; (8) a method of ameliorating
CC	symptoms associated with a condition mediated by a THAP/chemokine complex
CC	; (9) methods of identifying a test compound that modulates transcription
CC	at a THAP responsive element or that modulates the transport of a
CC	chemokine into the nucleus; (10) methods for reducing the symptoms
CC	associated with a condition selected from excessive or insufficient
CC	angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
CC	or insufficient apoptosis, cardiovascular disease and neurodegenerative
CC	diseases; symptoms associated with a condition resulting from the
CC	activity of a chemokine or a THAP-family polypeptide in an individual; or
CC	symptoms associated with transcriptional repression or activation
CC	mediated by a THAP-family polypeptide in an individual; (11) a vector
CC	comprising a THAP responsive promoter operably linked to a nucleic acid
CC	encoding a detectable product; (12) a genetically engineered cell
CC	comprising the vector described above or that expresses a THAP-family
CC	polypeptide or its biological fragment; (13) an in vitro transcription
CC	reaction comprising a nucleic acid comprising a THAP responsive promoter,
CC	ribonucleotides and an RNA polymerase; and (14) an isolated mutant THAP-
CC	family polypeptide that does not bind to a chemokine. The pharmaceutical
CC	composition has antiangiogenic, antiinflammatory, cardiovascular,
CC	cytostatic, neuroprotective and osteopathic activities, and can be used
CC	as a THAP and THAP synthesis modulator. The composition can be used for

CC modulating the expression of a THAP responsive gene. Modulation us useful
CC for reducing symptoms of conditions such as excessive or insufficient
CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
CC or insufficient apoptosis, cardiovascular disease or neurodegenerative
CC diseases. The present sequence is used in the exemplification of the
CC present invention.

XX
XX Sequence 371 AA;

Alignment Scores:
Pred. No.: 1.84e-194 Length: 371
Score: 1929.00 Matches: 364
Percent Similarity: 98.11% Conservative: 0
Best Local Similarity: 98.11% Mismatches: 7
Query Match: 98.32% Indels: 0
DB: 8 Gaps: 0

US-10-046-935-2234 (1-1116) x ADQ09178 (1-371)

QY 1 ATGACGCTCCGCGTCCGAGAAAGATCTCAGAGTAAAGAAACATTAAGAAATTC 60
Db 1 MetAspAlaArgValProGlnLysAspLeuArgValLysLeuLysAsnLeuLysPhe 20
QY 61 AGATATGTGAAGTTGAAATTCATTCATGGAAACCTCGTCATCTCTGATGACAGTGTGCACAGC 120
Db 21 ArgTyrValLysLeuLysSerMetGluThrSerSerSerSerSerSerSerSerSerSerSer 40
QY 121 TTTTGCTTCTGATAATTTTGCACACAGAGCTGCAGTCAGTTCGGGAGAGCTGTAGGACC 180
Db 41 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 60
QY 181 CGCAGCCAGTCAGGACACTCTGGACCTCTCAGGTTGGCGATCAAGTTTCCACGCGGAGT 240
Db 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80
QY 241 ACCAGGGGAGCAACCAACAAAAAGCAGAGTCCCGCCAGCCCTCAGAGAAATTTCTGTGACT 300
Db 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100
QY 301 GATTCCAACTCCGATTCAGAGATGAAAGTGAATGAAATTTTTGGAGAAAAGGGCTTTA 360
Db 101 AspSerAsnSerAspSerGluAspLysSerGlyMetAsnPheLeuGluLysArgAlaLeu 120
QY 361 AATATAAGCAAAACAAGCAATGCTTCACAAACTCATGTCTGTAATTAGAAAGCTTCCCT 420
Db 121 AsnLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140
QY 421 GGCTCGTTTCGGTGGAAACATCCCTCCAGGCTCCGACTCACAATCAAGGAGACCGCA 480
Db 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160
QY 481 AGCGGTACATTCGGGTGTGCTTCAGAGAAACCTCGAACGGAGAGCTCGTCTCTTT 540
Db 161 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgAlaArgProLeu 180
QY 541 ACCAGGTCAGGTCGCCGATCCTCGGGTCCCTTCGACGCTTACCCATCGRGAGAGCCAGT 600
Db 181 ThrArgSerArgSerArgLysLeuGlySerLeuAspAlaLeuProMetGluGluGlu 200
QY 601 TTTGARLWTWGTACATGTTGGTCAGAAAGAGGAGACCGTGGATGGCTACATGAATCAA 660
Db 201 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220
QY 661 GATGACTCGCCAGAACCGTCTCGTCAGATCATCCGTGACCCCTTCGCATATAATTCGC 720
Db 221 AspAspLeuProArgSerArgArgSerArgSerValThrLeuProHisIleIleArg 240
QY 721 CCAGTGGAGAAGATTACAGAGGAGAGCTTGGAGAACGCTCTCCAGCAATTTCTCGAGAGAG 780
Db 241 ProValGluGluIleThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLys 260
QY 781 ATATATAACCGTTCCACTGGGCTCTACTTGTTCATCAATGCCGTGACAGACCTATTGATACC 840

Db 261 IleTyrAsnArgSerIeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 280
 QY 841 AAAACAACTGCAGAAACCCAGACTGCTGGGGGTTCGAGGCCAGTTCGTGGCCCTGC 900
 Db 281 LysThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCys 300
 QY 901 CTTGCAAAACCGTTATGGTGAAGAGGTCAAGGATGCTCTCTGATCCGAACCTGGCATTCG 960
 Db 301 LeuArgAsnArgTyrGlyGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 320
 QY 961 CGCCTTGTGAGGAATCTGCAACTGCTTCTGCGCAGGAGATGCGAGGTGCG 1020
 Db 321 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 340
 QY 1021 ACTGGGCTCCTGTGATTTAGCAATATCATGGCTTTGGGATGTCATGCTACTTG 1080
 Db 341 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 360
 QY 1081 AAAAGCCTGAACAGCAATTTGAATGCAAGCA 1113
 Db 361 LysSerLeuLysGlnGluPheGluMetGlnAla 371

RESULT 7

AAB43720

ID AAB43720 standard; protein; 407 AA.

XX AAB43720;

DT 08-FEB-2001 (first entry)

XX Human cancer associated protein sequence SEQ ID NO:1165.

XX Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
 KW antidiabetic; antiaslthmatic; antirheumatic; antiarthritic; antiviral;
 KW dermatoflammatory; antithyroid; antiallergic; antibacterial; cardiant;
 KW vasotropic; antiporiatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening.

XX Homo sapiens.

XX OS

XX PN

XX WC2000055350-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000MO-US0005882.

XX 12-MAR-1999; 99US-0124270P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX PA

XX PI

XX DR

XX WPI; 2000-587533/55.

XX N-PSDB; AAC77929.

XX Novel isolated nucleic acids comprising sequences encoding peptides
 useful for treating or diagnosing e.g. cancer.

XX Claim 11; Page 1785-1787; 2352pp; English.

XX AAC77607 to AAC78448 encode the human cancer associated proteins given in
 CC ABA43398 to ABA44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnery; immunomodulator;
 CC antidiabetic; antiaslthmatic; antirheumatic; antiarthritic;
 CC antiinflammator; antithyroid; antiallergic; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
 CC vasotropic; antiporiatic and antiangiogenic. The

CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and ABA44240 represent sequences used in the exemplification of
 CC the present invention

XX Sequence 407 AA;

Alignment Scores:
 Pred. No.: 1 92e-194 Length: 407
 Score: 1929.00 Matches: 364
 Percent Similarity: 98.11% Conservative: 0
 Best Local Similarity: 98.11% Mismatches: 7
 Query Match: 98.32% Indels: 0
 DB: 3 Gaps: 0

US-10-046-935-2234 (1-1116) x AAB43720 (1-407)

QY 1 ATGACGCTCGCGCTGCCGAGAAAGATCTCAGATAAAGAACTTAAAGAATTC 60

Db 37 MetAspAlaArgArgValProGlnLysAspLeuArgValLysAsnLeuLysPhe 56

QY 61 AGATATGTGAAGTTGATTTCCATGGAAACCTCGTCATCTCTCATCAGACATCTGACAGC 120

Db 57 ArgTyrValLysLeuIleSerMetGluThrSerSerSerSerSerSerSerSerSer 76

QY 121 TTGTCTTCTGATAATTTTGCACACAGAGCTCGCTCAGTTCGGAAGGCTGTAGAAC 180

Db 77 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 96

QY 181 CGCAGCAGTGCAGGCACTCTGCACTCTCAGGCTGGCGATGAAGTTTCAGCGCGAGT 240

Db 97 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 116

QY 241 ACCAGGGAGCAACCAACAAAGAGAGTCCCGCCAGCCCTCAGAGATTTCTGTGACT 300

Db 117 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 136

QY 301 GATTCCAACCTCCGATTCAGAAAGATGAAAGTGAATGAAATTTTGGAGAAAAGGCTTTA 360

Db 137 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 156

QY 361 AATATAAGCAAAACAAAGCAATGCTTGCAAAACTCATGTCTGAAATAGAAAGTTCCT 420

Db 157 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 176

QY 421 GGCTCGTCCGTGGAGACATCCCTCCAGGCTCCGACTCACAATCAAGAGACCCGGA 480

Db 177 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 196

QY 481 AGCGGTACATTCCTCCGGGTGTGCTTCAGAGAGAAACCTCAACCGAGAGCTCGTCTCTT 540

Db 197 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgArgAlaArgProLeu 216

QY 541 ACCAGGTCAAGTCCCGGATTCCTCGGTCCTTTCAGCTCTACCCATGCGAGAGCCAGT 600

Db 217 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGlu 236

QY 601 TTGARKMTWMTGATCATGTTGGTGAAGAGAGAACCGTGGATGCTACATGAATGAA 660

Db 237 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 256

QY 661 GATGACCTGCCAGAGCCCTCGCTCCAGATCATCCGTGACCCCTCCGCAATATAATTCGC 720

Db 257 AspAspLeuProArgSerArgArgSerValThrLeuProHisIleLeuArg 276

QY 721 CCAGTGAAGAAATTCAGAGGAGGAGTTCGAGAACGTCGACGAATTCCTCGAGAGAAG 780
DB 277 ProValGluGluThrGluGluGluGluValCysSerAsnSerArgGluLys 296
QY 781 ATATATACCGTTCACGCGGCTCTACTGTCATCAATGCGCTCAGAGACTATTGATACC 840
DB 297 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 316
QY 841 AAAACAACTGCAGAACCCAGACTGCGGGCGTTCGAGGCCAGTCTGTGGCCCTGCG 900
DB 317 LysThrAsnCysArgAsnProAspCysTyrGlyValArgGlyGlnPheCysGlyProCys 336
QY 901 CTTGGAACCGTTATGCTGAAGAGGTGAGGATGCTGCTGGATCGAACTGCAATTGC 960
DB 337 LeuArgAsnArgTyrGlyGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 356
QY 961 CCGCTTGTGAGGAATCTCAACTGCACTGCTTTCGCGGAGGAGATGACGCTGTGCG 1020
DB 357 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 376
QY 1021 ACTGGGTCCTCTGTATTTAGCCAAATATCATGCTTGGGATGTGCATGCTACTTG 1080
DB 377 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 396
QY 1081 AAAAGCTGAACAGGAATTTGAAATGCAAGCA 1113
DB 397 LysSerLeuLysGlnGluPheGluMetGlnAla 407

RESULT 8

AAG74880

ID AAG74880 standard; protein; 407 AA.

AC AAG74880;

XX 03-SEP-2001 (first entry)

XX Human colon cancer antigen protein SEQ ID NO:5644.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;

XX colorectal carcinoma.

XX Homo sapiens.

XX WO200122920-A2.

XX 05-APR-2001.

XX 28-SEP-2000; 2000WO-US026524.

XX 29-SEP-1999; 99US-0157137P.

XX 03-NOV-1999; 99US-0163280P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Barash SC, Birse CE, Rosen CA;

XX WPI; 2001-235357/24.

XX N-PSDB; AAH34285.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,

XX useful for preventing, diagnosing and/or treating colorectal cancers.

XX Claim 11; Page 7194-7195; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon

XX cancer-associated nucleic acid molecules (N) and proteins (P), where the

XX proteins are collectively known as colon cancer antigens. The colon

XX cancer antigens have cytostatic activity and can be used in gene therapy

XX and vaccine production. N and P may be used in the prevention, diagnosis

XX and treatment of diseases associated with inappropriate P expression. For

XX example, N and P may be used to treat disorders associated with decreased

XX expression by rectifying mutations or deletions in a patient's genome

CC that affect the activity of P by expressing inactive proteins or to
CC supplement the patients own production of P. Additionally, N may be used
CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
CC into a host cell and culturing the cell to express the proteins. N and P
CC can be used in the prevention, diagnosis and treatment of colorectal
CC carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent
CC sequences used in the exemplification of the present invention. N.B.
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
CC time of publication, meaning no sequences are present for SEQ ID NO:1027
CC to 1052, 7921 and 7922
XX

SQ Sequence 407 AA;

Alignment Scores:

Pred. No.: 1,92e-194 Length: 407
Score: 1929.00 Matches: 364
Percent Similarity: 98.11% Conservative: 0
Best Local Similarity: 98.11% Mismatches: 7
Query Match: 98.32% Indels: 0
DB: 4 Gaps: 0

US-10-046-935-2234 (1-1116) x AAG74880 (1-407)

QY 1 ATGGAGCGCTCGCGCGTGGCGAGAAAGATCTCAGAGTAAAGAAAGAACTTAAAGAAATTC 60
DB 37 MetAspAlaArgValProGlnLysAspLeuArgValLysAsnLeuLysLysPhe 56
QY 61 AGATATGTGAAGTTGATTTCCATGGAACCTGTCATCTCTCTGATGACAGTTGTGACAGC 120
DB 57 ArgTyrValLysLeuIleSerMetGluThrSerSerSerSerSerSerSerSerSer 76
QY 121 TTTGCTCTTCATATTTTCGAAACACAGAGCTCGAGTCAGTTCGGAGAGCTGTAGAAC 180
DB 77 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 96
QY 181 CGCAGCAGTGCAGGCACTCTGGACCTCTCAGGCTGGCGATGAAGTTTTCAGCGCGAGT 240
DB 97 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 116
QY 241 ACCAGGGGACCAACCAACAAAAGCAGAGTCCCGCAGCCCTCAGAGAATTCGTGACT 300
DB 117 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 136
QY 301 GATTCCAACTCCGATTCAGAGATGAAGTGGATGAATTTTGGAGAAAAGGGCTTTA 360
DB 137 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 156
QY 361 AATATAAGCAAAACAAAGCAATGCTTGCAAACTCATGTCTGAATTAGAAAGCTTCCCT 420
DB 157 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 176
QY 421 GGCTCGTTCGGTGAAGACATCCCTCCAGGCTCCGACTCACAATCAAGAGACCGCGA 480
DB 177 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 196
QY 481 AGCGCTACATTCCTCGGCTGCTTCCAGAGAAACCTGACCGAGAGCTGTCTCTCTT 540
DB 197 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgAlaArgProLeu 216
QY 541 ACCAGTCAAGTCCCGGATCCTCGGCTCCTTGACGCTCTACCCATGCGAGAGCACT 600
DB 217 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGlu 236
QY 601 TTKARKMTMWGTACATGTTGTGAGAAAGAGACCGTCGATGGCTGCTACATGATGAA 660
DB 237 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 256
QY 661 GATGACCTGCCAGAGCGCTCGCTCCAGATCATCGGTGACCCCTCCCGCATATATTCG 720
DB 257 AspAspLeuProArgSerArgArgSerArgSerValThrLeuProHisIleLeArg 276
QY 721 CCAGTGGAGAAATTTACAGAGGAGGAGTGGAGAACGCTCTGAGCAATTTCTCGAGAGAAG 780

Db 277 ProValGluGluIleThrGluGluGluLeuAenValCysSerAsnSerArgGluLys 296
QY 781 ATATATAACCGTTCACTGGCTTACTTGTCTATCAATCCGCTCAGAACACTATTGATACC 840
Db 297 IletyAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 316
QY 841 AAAACAAACTGCAGAAACCCAGACTGCTGGGGCGTTTCGAGGCCAGTTCTGTGGCCCTGC 900
Db 317 LysThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCys 336
QY 901 CTTGAAACCGTTATGGTGAAGAGTTCAGGGATGCTCTGCTGGATCCGAACTGGCATTCG 960
Db 337 LeuArgAsnArgTyrrGlyGluGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 356
QY 961 CGCCTTGTTCGAGGAATCTGCAACTGCAGTTTCTCGGCGCAGCAGATGAGCGTGGG 1020
Db 357 ProProCysArgGlyCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 376
QY 1021 ACTGGGGTCTTGTGTATTTAGCCAAATATCATGGCTTTTGGGAATGTGCATCGCTACTTG 1080
Db 377 ThrGlyValLeuValTyrrLeuAlaLysTyrrHisGlyPheGlyAsnValHisAlaTyrrLeu 396
QY 1081 AAAAGCTGAACACAGAAATTCGAATGCAAGCA 1113
Db 397 LysSerLeuLysGlnGluPheGluMetGlnAla 407

RESULT 9

ADA98284

ID ADA98284 standard; protein; 397 AA.

XX AC

ADA98284;

XX DT

20-NOV-2003 (first entry)

XX DE

Human secreted protein sequence #125.

XX KW

human; secreted protein; cardiovascular disorder; arrhythmia;
atherosclerosis; stroke; endocarditis; congestive heart failure;
rheumatic heart disease; cardiomyopathy; haemorrhoids; varicose veins;
migraine; thrombosis; neural disorder; immune system disorder;
muscular disorder; reproductive disorder; gastrointestinal disorder;
pulmonary disorder; renal disorder; proliferative disorder; cancer.

XX OS

Homo sapiens.

XX XX

WQ2003004623-A2.

XX PD

16-JAN-2003.

XX XX

26-MAR-2002; 2002WO-US0009922.

XX PF

27-MAR-2001; 2001US-0278650P.

XX PR

12-SEP-2001; 2001US-00950082.

XX PR

12-SEP-2001; 2001US-00950083.

XX XX

(HUMA-) HUMAN GENOME SCI INC.

XX PA

Rosen CA, Ruben SM;

XX PI

WPI; 2003-247946/24.

XX DR

New human secreted polypeptide and nucleic acid molecules, useful for
diagnosing, preventing, prognosticating or treating cardiovascular
disorders (e.g. arrhythmia, atherosclerosis, cardiomyopathy, or
thrombosis).

XX PT

Claim 1; SEQ ID NO 392; 1572pp; English.

XX PS

The invention comprises the amino acid and coding sequence of human
secreted proteins. The DNA and protein sequences of the invention are
useful in the treatment of cardiovascular disorders, such as: arrhythmia,
atherosclerosis, stroke, endocarditis, congestive heart failure,
rheumatic heart disease, cardiomyopathy, haemorrhoids, varicose veins,

CC migraine, or thrombosis. The DNA and protein sequences may also be used
for treating or preventing: neural disorders, immune system disorders,
CC muscular disorders, reproductive disorders, gastrointestinal disorders,
CC pulmonary disorders, renal disorders, proliferative disorders and/or
CC cancerous diseases. The present amino acid sequence represents a human
secreted protein of the invention. NOTE: The present sequence is shown on
CC the WFO website.

XX

SQ Sequence 397 AA;

Alignment Scores:

Pred. No.: 6,42e-194 Length: 397
Score: 1924.00 Matches: 364
Percent Similarity: 97.85% Conservative: 0
Best Local Similarity: 97.85% Mismatches: 7
Query Match: 98.06% Indels: 1
DB: Gaps: 0

US-10-046-935-2234 (1-1116) x ADA98284 (1-397)

QY 1 ATGGAGCGTCCGCGGTGCGGAGAAAGATCTCAGAGTAAAGAACTTAAGAATTC 60

Db 1 MetAspAlaArgValProGlnLysAspLeuArgValLysLysAsnLeuLysLysPhe 20

QY 61 AGATATGTGAAGTTGATTTCCATGAAACCTCGTCTCTCTGATCAGAGTTGTGACAG 120

Db 21 ArgTyrrValLysLeuIleSerMetGluThrSerSerSerSerSerSerSerSerSer 40

QY 121 TTTTGTCTCTGATAATTTTGCACACGAGGCTGCAGTTCAGTTCCGGGAAGGCTGTAGAGC 180

Db 41 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGlyCysArgThr 60

QY 181 CGCAGCAGTGCAGGCACCTCTCGACCTCTCAGGGTGGCGATGAAGTTTCAGCGCGGAGT 240

Db 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80

QY 241 ACAGGGAGCAACCAACAAAAGAGAGTCCGCGACCTCAGAGATTCGTGACT 300

Db 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100

QY 301 GATTCACATCCGATTTCAGAGTGAAGTGAATGAATTTTGGAGAGAAAGGGCTTTA 360

Db 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 120

QY 361 AATATAAGCAAAACAAACCAATGTTTGCACCACTCATGTCTGAATAGAAAGCTTCCCT 420

Db 121 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140

QY 421 GGCTCGTTCCGTGGAAGACATCCCTCCAGGCTCCGACTCACAATCAAGGACCGCGA 480

Db 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160

QY 481 AGCGGTACATTCGCGGTGTTGCTTCCAGAGAAACCTGAAACGAGAGCTCGTCTCTT 540

Db 161 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgArgAlaArgProLeu 180

QY 541 ACCAGGTCAAGTCCCGGATTCCTCGGTCCTTTCAGCTTACCCATTCGAGAGAGCCAGT 600

Db 181 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGlu 200

QY 601 TTKGARKMTWGTACATGTTGGTGAAGAGAGACCGTGGATGGCTACATGAATGAA 660

Db 201 GluGluAspLysTyrrMetLeuValArgLysArgLysThrValAspGlyTyrrMetAsnGlu 220

QY 661 GATGACTGCCAGAGAGCGCTCGCTCCAGATCATCGTGCACCCCTCCGCAATATAATTCGC 720

Db 221 AspAspLeuProArgSerArgArgSerArgSerValThrLeuProHisIleIleArg 240

QY 721 CCAGTGAAGAAATTCAGAGAGAGAGTGGAGAACGTCTGCAGCAATTCCTCGAGAGAG 780

Db 241 ProValGluGluIleThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLys 260

QY 781 ATATATAACCGTTCACTGGGCTCTACTTGTTCATCAATGCGCTCAGAGAGACTATTGATACC 840

```
Db 261 IIEYASNAAGSERLEUGLYSERTHRYSHISGLNCYSARGGLNLYSTHRIEASPTH 280
QY 841 AAAACAAACTGCAGAAACCCAGACTGCTGGGCGCTTCGAGGCCAGTCTCTGTGGCCCTGC 900
Db 281 LYSSTRASNCYSARGASNPASPCYSITPGLYVALARGGLYNPHECYSGLYPROCY 300
QY 901 CTTGGAACCGTTATGTTGAGAGAGTCAGGGATGCTCTGCTGGATCCGAACCTGGCATTGC 960
Db 301 LEUARGASNAAGTYRGLYGLUVALARGASPLAALAULEUASPASNTTRPHISCYS 320
QY 961 CGCCTCTGTCGAGGAATCTGCAACTGCAGTTCTCCGCGCAGCAGATGAGCGGTGTGCG 1020
Db 321 PROPCYASARGGLYLECYSASNCYSERPHECYSARGGLNARGASPLAALAULEUASP 340
QY 1021 ACTGGGTCCTTGTGTTATTTAGCAAAATATCATGCTTTTGGAAATGTCATGCCCTACTTG 1080
Db 341 THRGLYVALLEUVALTYRLEUALALYSTYRHSGLYPHEGLYASNVALHISALATYRLEU 360
QY 1081 AAAAGCCTGAACAGGAATTTGAAATGCAAGCAT 1114
Db 361 LYSERLEU-ASNARGASNLEULYSYSLYS 371

RESULT 10
ADA44119
ID ADA44119 standard; protein; 397 AA.
XX
AC ADA44119;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human secreted protein SEQ ID 311.
XX
KW Gene therapy; human; Antidiabetic; Anorectic; Ophthalmological;
KW Neuroprotective; Cerebroprotective; Antianemic.
XX
OS Homo sapiens.
XX
PN W0203000865-A2.
XX
PD 03-JAN-2003.
XX
PF 26-MAR-2002; 2002WO-US009105.
XX
PR 27-MAR-2001; 2001US-0278650P.
PR 12-SEP-2001; 2001US-00950082.
PR 12-SEP-2001; 2001US-00950083.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2003-184045/18.
DR N-PSDB; ADA43929.
XX
PT A human secreted protein and nucleic acids useful for preparing a
PT diagnostic or pharmaceutical composition for diagnosing or treating
PT diabetes or conditions related to diabetes, e.g. hyperglycemia, obesity,
PT retinopathy, neuropathy.
XX
PS Claim 1; SEQ ID NO 311; 701pp; English.
XX
CC The invention relates to novel genes and their fragments which are useful
CC for preventing, treating or ameliorating medical conditions e.g. by
CC protein or gene therapy. The genes are isolated from a range of human
CC tissues disclosed in the specification. The nucleic acids and proteins
CC are useful in the diagnosis, treatment and prevention of conditions
CC related to diabetes, e.g. hyperglycaemia, obesity, retinopathy,
CC polynuropathy, atherosclerosis, anaemia, stroke, gangrene, impotence,
CC infection, cataract, renal disorders, or endocrine disorders. The present
CC sequence was used to illustrate the invention.
XX
SQ Sequence 397 AA;
```

```
Alignment Scores:
Pred. No.: 6,428-194 Length: 397
Score: 1924.00 Matches: 364
Percent Similarity: 97.85% Conservative: 0
Best Local Similarity: 97.85% Mismatches: 7
Query Match: 98.06% Indels: 1
DB: 6 Gaps: 0

US-10-046-935-2234 (1-1116) x ADA44119 (1-397)
QY 1 ATGACGCTCCGCGCTGCCAGAAAGATCTCAGAGTAAGAAGAACTTAAGAAATTC 60
Db 1 MetAspAlaAaArgValProGlnLysAspLeuArgValLysLysAsnLeuLysPhe 20
QY 61 AGATATGTAAGTGAATTTCCATTCGAAACCTCGTCATCTCTGATGACAGTTGTGACAGC 120
Db 21 ArgTyrValLysLeuLysSerMetGluThrSerSerSerSerSerSerSerSerSerSer 40
QY 121 TTTGCTTCTGATTAATTTTGCACACGAGGCTGAGTCAGTTCGGGAAAGGCTGTAGGACC 180
Db 41 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 60
QY 181 CGCAGCCAGTGCAGGCATCTGACCTCTCAGGCTGCGATGAGTAACTTTCCAGCGCGGACT 240
Db 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80
QY 241 ACAGGGGAGCAACAAACAAAGAGAGTCCCGCCAGCCCTCAGAGAAATTTCTGTGACT 300
Db 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100
QY 301 GATTCCAACTCCGATTTCAGAGATGAAGTGAATGAATTTTGGAGAGAAAGGCTTTA 360
Db 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 120
QY 361 AATATAAGCAAAACAAAGCAATGCTTGCAAACTCATGTCTGAATAGAAAGCTTCCT 420
Db 121 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140
QY 421 GGCTGCTGCTGGAGAGACATCCCTCCAGGCTCCGACTCACAATCAAGAGAGACCCGA 480
Db 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160
QY 481 AGCGTACATTCGCGGTGTTGCTTCAGAGAGAAACCTGAAAGAGAGTGGTCTCTT 540
Db 161 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgArgAlaArgProLeu 180
QY 541 ACCAGGTCAAGGTCCCGGATCCTCGGTCCTTTCAGCTCTACCCATGGRGAGGCCAGT 600
Db 181 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGlu 200
QY 601 TTGARKMTWVGATATGTTGGTGAAGAGAGAGCCGTCGATGGATGCTACATGAATGAA 660
Db 201 GluGluAspLysTyrMetLeuValArgLysThrValAspGlyTyrMetAsnGlu 220
QY 661 GATGACCTCCAGAGACCGTCTCCAGATCATCGTGACCTTCGCGATATAATTCGC 720
Db 221 AspAspLeuProArgSerArgArgSerArgSerValThrLeuProHisIleIleArg 240
QY 721 CCAGTGGAAAGAAATTAACAGAGAGAGTGGAGAGCTGCGAGCAATTCGAGAGAG 780
Db 241 ProValGluGluIleThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLys 260
QY 781 ATATATAACCGTTCAGTGGCTCTACTTGTCTCAATGCCGTGAGAGAACTATTGATACC 840
Db 261 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 280
QY 841 AAAACAACTGCAGAAACCCAGACTGCTGGGCGCTTCGAGGCCAGTCTCTGTGGCCCTGC 900
Db 281 LYSSTRASNCYSARGASNPASPCYSITPGLYVALARGGLYNPHECYSGLYPROCY 300
QY 901 CTTGGAACCGTTATGTTGAGAGAGTCAGGGATGCTCTGCTGGATCCGAACCTGGCATTGC 960
```


Db 301 LeuArgAsnArgTyrGlyGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 320
 QY 961 CGCCTTTCGAGGAATCTGCACTGCTTCTCCGCGCAGAGATGAGCGGTGCG 1020
 Db 321 ProProCysArgGlyLeuCysAenCysSerPheCysArgGlnArgAspGlyArgCysAla 340
 QY 1021 ACTGGGCTCTGTGATTTAGCAATATCATGCTTTGGGAATGTGATCGCTACTTG 1080
 Db 341 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 360
 QY 1081 AAAAGCCTGAAACAGGAATTTGAAATGCAAGCAT 1114
 Db 361 LysSerLeu-AsnArgAsnLeuLysCysLysHis 371
 RESULT 11
 ADC20449
 ID ADC20449 standard; protein; 397 AA.
 XX
 AC ADC20449;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Human secreted protein - amino acid sequence #130.
 XX
 KW gene therapy; human; secreted protein; haemopoietic disorder;
 KW haematological disorder; anaemia; haemophilia; inflammatory disorder;
 KW inflammatory bowel disease; Crohn's disease; neoplastic disease; cancer;
 KW leukemia; wound healing; epithelial cell proliferation disorder;
 KW immune disorder; autoimmune disorder; asthmatic disorder;
 KW cardiovascular disorder; atherosclerosis; myocarditis;
 KW infectious disease; HIV; AIDS; endocrine disorder; diabetes;
 KW gastrointestinal disorder; duodenal ulcer; gastroenteritis.
 XX
 OS Homo sapiens.
 OS
 PN WC020292787-A2.
 XX
 PD 21-NOV-2002.
 XX
 PF 26-MAR-2002; 2002WO-US009257.
 XX
 PR 27-MAR-2001; 2001US-0278650P.
 PR 12-SEP-2001; 2001US-00950082.
 PR 12-SEP-2001; 2001US-00950083.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2003-123287/12.
 XX
 PT New human secreted proteins and nucleic acid molecules, useful for
 PT preparing a diagnostic or pharmaceutical composition for diagnosing,
 PT preventing or treating hematopoietic or hematologic disorders, e.g.
 PT anemia or hemophilia.
 XX
 PS Claim 1; SEQ ID NO 403; 1512pp; English.
 XX
 CC The invention comprises the amino acid and coding sequences of human
 CC secreted proteins. The DNA and protein sequences of the invention are
 CC useful for detecting, preventing, diagnosing, prognosticating, treating
 CC or ameliorating: haematopoietic or haematological disorders (e.g. anaemia
 CC and hemophilia); inflammatory disorders (e.g. inflammatory bowel disease
 CC and Crohn's disease); neoplastic disease (e.g. cancer and leukaemia);
 CC wound healing and disorders of epithelial cell proliferation; immune
 CC disorders (e.g. autoimmune disorders and asthmatic disorders);
 CC cardiovascular disorders (e.g. atherosclerosis and myocarditis);
 CC infectious disease (e.g. HIV/AIDS); endocrine disorders (e.g. diabetes);
 CC and gastrointestinal disorders (e.g. duodenal ulcers and
 CC gastroenteritis). The present amino acid sequence represents a human
 XX secreted protein of the invention.
 XX
 SQ Sequence 397 AA;

Alignment Scores:
 Pred. No.: 6,42e-194 Length: 397
 Score: 1924.00 Matches: 364
 Percent Similarity: 97.85% Conservative: 0
 Best Local Similarity: 97.85% Mismatches: 7
 Query Match: 98.06% Indels: 1
 DB: 7 Gaps: 0
 US-10-046-935-2234 (1-1116) x ADC20449 (1-397)
 QY 1 ATGGAGCTCGCGGTGCGCGAGAAAGATCTCAGATGAAAGAAAGAACTTAAAGAAATTC 60
 Db 1 MetAspAlaArgValProGlnLysAspLeuArgValLysLysAsnLeuLysLysPhe 20
 QY 61 AGATATGTGAAGTTGATTTCCATGAAACCTCGTCACTCTCATGACAGTTGTGACGC 120
 Db 21 ArgTyrValLysLeuLysSerMetGluThrSerSerSerSerAspAspSerCysAspSer 40
 QY 121 TTTGCTTCTGATAATTTTGCACACGAGCGTGCAGTTCGCGAAGCGCTGTAGGACC 180
 Db 41 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 60
 QY 181 CGCAGCAGTGCAGGCACTCTGCACCTCTCAGGTCGCGATGAAGTTTCCAGCGCGGAGT 240
 Db 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80
 QY 241 ACCAGGCGAGCAACCAACAAAGAGAGTCCCGCCAGCCCTCAGAGAATTTCTGTGACT 300
 Db 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100
 QY 301 GATTCACATCCCATTCAGAGATGAAGTGAATGATTTTGGAGAAAGGCGCTTTA 360
 Db 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 120
 QY 361 AATATAAGCAAAACAAAGCAATGCTTGCAAACTCATGTCTCAATTAGAAGTTCCTCCT 420
 Db 121 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140
 QY 421 GGCTCGTTCCGTGGAAGACATCCCTCCAGGCTCCGACTCACAATCAAGAGACCCGGA 480
 Db 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160
 QY 481 AGCGTACATTCGCGGTGTTGCTTCCAGAGAAACCTGACCGAGAGCTCGTCTCTT 540
 Db 161 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgArgAlaArgProLeu 180
 QY 541 ACCAGTCAAGTCCCGGATCCCTCGGTCCTTACCGCTCTTACCCATGGRGAGAGCCAGT 600
 Db 181 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGlu 200
 QY 601 TTKGARKMTWGTACATGTTGGTGAGAAAGAGACCCGTTGATGCTACATGAATGAA 660
 Db 201 GluLysAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220
 QY 661 GATGACCTGCCAGAACCGCTCGCTCCAGATCATCCGTGACCCCTCCGATATATTCGC 720
 Db 221 AspAspLeuProArgSerArgSerValThrLeuProHisIleLeuArg 240
 QY 721 CCAGTGAAGAAATTTACAGAGGAGGAGTTGAGAAACGTTCTGCAGCAATTTCTCGAGAGAG 780
 Db 241 ProValGluGluIleThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLys 260
 QY 781 ATATATAACCGTTCACTGGGCTCTACTTGTGTCATCATCCCGCTCAGAACACTATTGATAC 840
 Db 261 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 280
 QY 841 AAAACAACTCCAGAAACCCAGACTCTGGGGCGGTTCGAGGCGAGTTCGTGGCCCGCTGC 900
 Db 281 LysThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCys 300
 QY 901 CTTTCGAAAACCGTTATGTTGTAAGAGGTCAGGGATGCTCTGCTGGATCCGAACTGGCATTC 960

Db 301 LeuArgAsnArgTyrGlyGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 320
 Qy 961 CCGCTTGTTCGAGGAATCTGCAACTGAGTTCTCGCGCAGCAGAGATGAGCGGTGTCG 1020
 Db 321 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 340
 Qy 1021 ACTGGGCTCCTTGTGTATTATAGCAATATCATGCTTTGGAAATGTGCATGCCCTACTTG 1080
 Db 341 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 360
 Qy 1081 AAAAGCCTGAAACAGGAATTGAAATGCAAGCAT 1114
 Db 361 LysSerLeu-AsnArgAsnLeuLysCysLysHis 371

RESULT 12

ADFI0748

ID ADFI0748 standard; protein; 397 AA.

XX AC ADFI0748;

XX AC ADFI0748;

XX AC ADFI0748;

DT 12-FEB-2004 (first entry)

XX Human secreted protein #70.

XX H6EDM64; HBHAA05; HBJCR46; HBKDL16; HCMXS51; HCQBH72; HDPPQ30; HE2CM39;

XX HE2EAL0; HGBHP91; HLDQU79; Cytostatic; Hepatotropic; Antidiabetic;

XX Antiinflammatory; neuroprotective; Anti-HIV; Vulnerary; Gynecological;

XX Antiinfertility; Gene therapy; gastrointestinal disorder; cancer;

XX Alzheimer's disease; chromosome identification.

XX Homo sapiens.

XX OS

XX PN

XX WO200299085-A2.

XX 12-DEC-2002.

XX 26-MAR-2002; 2002WO-US009135.

XX 27-MAR-2001; 2001US-0278650P.

XX 12-SEP-2001; 2001US-00950082.

XX 12-SEP-2001; 2001US-00950083.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2003-221310/21.

XX New human secreted polypeptides for diagnosing and treating neural,

XX immune system, muscular, reproductive, gastrointestinal, cardiovascular,

XX renal, and proliferative disorders and cancerous diseases.

XX Claim 1; SEQ ID NO 211; 855pp; English.

XX The present invention relates to an isolated polypeptide chosen from 123

XX human secreted proteins, such as, H6EDM64, HBHAA05, HBJCR46, HBKDL16,

XX HCMXS51, HCQBH72, HDPPQ30, HE2CM39, HB9EA10, HGBHP91 and HLDQU79. The

XX polypeptides are useful for the preparation of a diagnostic or

XX pharmaceutical composition for diagnosing or and are useful for treating

XX or preventing diseases or conditions, such as neural, immune system,

XX muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,

XX renal, proliferative disorders and cancerous diseases and conditions. The

XX polypeptides have immune activity, chemotactic activity, and binding

XX activity to treat and prevent neuronal damage which occurs in certain

XX neuronal disorders or neuro-degenerative conditions such as Alzheimer's

XX disease, Parkinson's disease, and acquired immunodeficiency syndrome

XX (AIDS)-related complex, and to prevent skin aging due to sunburn by

XX stimulating keratinocyte growth. The molecules are also useful to

XX modulate mammalian characteristics including .The encoding sequences are

XX useful for chromosome identification, radiation hybrid mapping, in gene

XX therapy, for identifying individuals from minute biological samples, as

XX additional DNA markers for restriction fragment length polymorphism

XX (RFLP), in forensic biology, molecular weight markers on Southern gels,

CC as diagnostic probes for the presence of a specific mRNA in a particular
 CC cell type, to raise anti-DNA antibodies using DNA immunization
 CC techniques, and as an antigen to elicit an immune response. The present
 CC sequence represents a human secreted protein of the invention.

XX

SQ Sequence 397 AA;

Alignment Scores:

Pred. No.: 6,42e-194 Length: 397
 Score: 1924.00 Matches: 364
 Percent Similarity: 97.85% Conservative: 0
 Best Local Similarity: 97.85% Mismatches: 7
 Query Match: 98.06% Indels: 1
 DB: 7 Gaps: 0

US-10-046-935-2234 (1-1116) x ADFI0748 (1-397)

Qy 1 ATGACGCTCGCGCGTGGCCAGAAAGATCTCAGAGTAAAGAAAGAACTTAAAGAAATTC 60

Db 1 MetAspAlaArgValProGlnLysAspLeuArgValLysLysAsnLeuLysLysPhe 20

Qy 61 AGATATGTGAAGTGTATTTCCATGGAAACCTCGTCATCTCTGATCACAGTTGTGACAGC 120

Db 21 ArgTyrValLysLeuIleSerMetGluThrSerSerSerSerSerSerSerSerSerSer 40

Qy 121 TTTGCTTCTCATTAATTTTGCACACAGAGCTCGAGTCACTTCGGGAAGGCTGTAGGACC 180

Db 41 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 60

Qy 181 CGCAGCAGTGCAGGCACTCTGCACCTCTCAGGCTGGCGATGAAGTTTCCAGCGCGGAGT 240

Db 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80

Qy 241 ACAGGGAGCAACCAACAAAAAGCAGAGTCCCGCCAGCCCTCAGAGAATTTCTGTGACT 300

Db 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100

Qy 301 GATTCCAACCTCCGATTCAGAGATGAAGTGGAAATGAATTTTGGAGAAAAGGGCTTTA 360

Db 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 120

Qy 361 AATATAAGCAAAACAAAGCAATGCTTGCAAACTCATGTCTGAATTAGAAAGCTTCCCT 420

Db 121 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140

Qy 421 GGCTCGTTCCTGGGAAGACATCCCTCCAGGCTCCGACTCAACATCAAGGAGACCGCGA 480

Db 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160

Qy 481 AGCGGTACATTCGCGGTGCTCTCCAGGAGAAACCTTGAACGGAGAGCTGCTCTCTT 540

Db 161 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgAlaArgProLeu 180

Qy 541 ACCAGTCAAGTCCCGGATCCTCGGTCCTCGCTTACCTACCTACCTACCTACCTACCT 600

Db 181 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGlu 200

Qy 601 TTKGARKMTMGTCATGTTGGTGAAGAGAGACCGTGGATGGCTGCTACATGAATGAA 660

Db 201 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220

Qy 661 GATGACCTGCCAGAGCGGTCCGCTCCAGATCATCCGTGACCTTCCGATATATATTCG 720

Db 221 AspAspLeuProArgSerArgArgSerValThrLeuProHisIleArg 240

Qy 721 CCAGTGAAGAAATACAGAGAGAGTGGAGAACGCTGCAGCAATTCGAGAGAG 780

Db 241 ProValGluIleThrGluGluGluLeuAsnValCysSerAsnSerArgGluLys 260

Qy 781 ATATATAACCGTTCACTGGGTCTTCTTGTTCATCAATGCCGTTCAGAGACTATTGATACC 840

Db 261 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 280

QY 544 AGGTCAGGTCCTCGGATCTCCCTGACGCTCTACCCATGGRGAGAGCCAGTTTK 603
 |||||
 Db 261 ArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGluGlu 280
 |||||
 QY 604 GARKMTWGTACATGTTGGTGAGAAAGAGAGACCGTGGATGCTACATGAATGAAGAT 663
 |||||
 Db 281 GluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGluAsp 300
 |||||
 QY 664 GACCTGCCAGAACCGTCGCTCAGATCATCCGTGACCCCTCCGATATAATTCGCCCA 723
 |||||
 Db 301 AspLeuProArgSerArgArgSerArgSerValThrLeuProHisIleIleArgPro 320
 |||||
 QY 724 GTGGAAGAAATTACAGAGGAGGAGTTCGAGAACGCTCTCGACCAATTCCTCGAGAGAAGATA 783
 |||||
 Db 321 ValGluGluIleThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLysIle 340
 |||||
 QY 784 TATAACCGTTCACTGGGCTCTACTTGTATCAATGCGCGTGCAGAGACTATGTATACCAA 843
 |||||
 Db 341 TyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThrLys 360
 |||||
 QY 844 ACAACTGCAGAAACCCAGACTGCTGGGGGCTTCGAGGCCAGTCTGTGGCCCTGCGCTT 903
 |||||
 Db 361 ThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCysLeu 380
 |||||
 QY 904 CGAAACCGTTATGCTGAAGAGGTGAGGATGCTCTGCTGGATCCGAATCGCATTTGCCCG 963
 |||||
 Db 381 ArgAsnArgTyrGlyGluGluValArgAspAlaLeuLeuAspProAsnTrpHisCysPro 400
 |||||
 QY 964 CTTGTGAGGATCTGCAACTCGACTTCTGCGCGGAGGAGATGACCGGTGCGACT 1023
 |||||
 Db 401 ProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAlaThr 420
 |||||
 QY 1024 GGGGTCCTGTATTTAGCCAAATATCATGGCTTTGGGATGTCATGCTGCTACTTCAA 1083
 |||||
 Db 421 GlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeuLys 440
 |||||
 QY 1084 AGCTGAAACAGGAATTTGAAATGCAAGCA 1113
 |||||
 Db 441 SerLeuLysGlnGluPheGluMetGlnAla 450
 |||||

RESULT 14
 ID ADL91540
 AC ADL91540
 XX ADL91540;
 DT 17-JUN-2004 (first entry)
 DE Human immune-related polypeptide PRO69531, SEQ ID NO:55.

KW Human; PRO; activated T cell; immune-related; drug screening; detection;
 KW stimulation; immune response; stimulation; diagnosis; immune disorder;
 KW systemic lupus erythematosus; rheumatoid arthritis; immune deficiency;
 KW juvenile chronic arthritis; spondyloarthritis; osteoarthritis;
 KW idiopathic inflammatory myopathy; Sjogren's syndrome;
 KW systemic vasculitis; sarcoidosis; autoimmune haemolytic anaemia;
 KW autoimmune thrombocytopenia; thyroiditis; diabetes mellitus;
 KW immune-mediated renal disease; demyelinating disease;
 KW idiopathic demyelinating polynuropathy; Guillain-Barre syndrome;
 KW chronic inflammatory demyelinating polynuropathy; hepatobiliary disease;
 KW chronic active hepatitis; primary biliary cirrhosis;
 KW granulomatous hepatitis; sclerosing cholangitis;
 KW inflammatory bowel disease; gluten-sensitive enteropathy;
 KW Whipple's disease; autoimmune skin disease; immune-mediated skin disease;
 KW bullous skin disease; erythema multiforme; contact dermatitis; psoriasis;
 KW allergic disease; asthma; allergic rhinitis; atopic dermatitis;
 KW food hypersensitivity; urticaria; eosinophilic pneumonia;
 KW idiopathic pulmonary fibrosis; hypersensitivity pneumonitis;
 KW transplantation associated disease; graft rejection;
 KW graft-versus-host-disease; immunosuppressive; dermatological;
 KW hepatotropic; nephrotropic; antidiabetic; antiaesthetic; antipsoriatic;
 KW anti-allergic; antianaemic; antiarteriosclerotic; antiarthritic;

KW neuroprotective; respiratory; antiinflammatory; gene therapy.
 XX Homo sapiens.
 OS WO2004024072-A2.
 FN 25-MAR-2004.
 PD 10-SEP-2003; 2003WO-US028317.
 PF 11-SEP-2002; 2002US-0410340P.
 PR (GETH) GENENTECH INC.
 PA Bodary SC, Clark H, Hunte B, Jackman JK, Schoenfeld JR;
 PI Williams PM, Wood WI, Wu TD;
 PI WPI; 2004-329384/30.
 DR N-PSDB; ADL91539.

XX New PRO nucleic acid, useful for preparing a composition for diagnosing
 PT or treating an immune related disorder, e.g., systemic lupus
 PT erythematosus in a mammal.
 XX Claim 10; SEQ ID NO 55; 199pp; English.

CC The invention relates to isolated human immune-related polypeptides
 CC (designated PRO) and nucleic acids (ADL91486-ADL91587). The PRO
 CC polypeptides are overexpressed in CD4+ T cells activated by anti-CD3/ICAM
 CC -1 or anti-CD3/anti-CD28 antibodies compared with resting T cells and are
 CC useful as diagnostic markers and therapeutic targets for immune
 CC disorders. The invention also relates to sequences at least 80% identical
 CC to the PRO nucleic acid and polypeptide sequences of the invention;
 CC recombinant vectors and host cells comprising a PRO nucleic acid; a
 CC method for the recombinant production of a PRO polypeptide; antibodies
 CC against a PRO polypeptide; fusion proteins comprising a PRO polypeptide;
 CC methods of screening for compounds which modulate PRO polypeptide
 CC activity or expression; a method for detecting a PRO polypeptide; a
 CC method of detecting an immune response in a mammal; a method for
 CC stimulating an immune response in a mammal; and methods for diagnosing
 CC and treating immune-related disorders. PRO polypeptides and nucleic acids
 CC are useful in the diagnosis and treatment of immune-related disorders
 CC such as systemic lupus erythematosus, rheumatoid arthritis,
 CC osteoarthritis, juvenile chronic arthritis, spondyloarthritis, systemic
 CC sclerosis, idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
 CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
 CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
 CC disease, a demyelinating disease of the central or peripheral nervous
 CC system, idiopathic demyelinating polynuropathy, Guillain-Barre syndrome,
 CC chronic inflammatory demyelinating polynuropathy, hepatobiliary disease,
 CC infectious or autoimmune chronic active hepatitis, primary biliary
 CC cirrhosis, granulomatous hepatitis, sclerosing cholangitis, inflammatory
 CC bowel disease, gluten-sensitive enteropathy, Whipple's disease,
 CC autoimmune or immune-mediated skin disease, bullous skin disease,
 CC erythema multiforme, contact dermatitis, psoriasis, allergic disease,
 CC asthma, allergic rhinitis, atopic dermatitis, food hypersensitivity,
 CC urticaria, immunologic disease of the lung, eosinophilic pneumonitis,
 CC idiopathic pulmonary fibrosis, hypersensitivity pneumonitis,
 CC transplantation associated disease, graft rejection or graft-versus-host-
 CC disease. The present sequence represents a human immune-related PRO
 CC polypeptide of the invention.

XX Sequence 450 AA;
 Alignment Scores:
 Pred. No.: 3,49e-189 Length: 450
 Score: 1879.50 Matches: 364
 Percent Similarity: 80.8% Conservative: 0
 Best Local Similarity: 80.8% Mismatches: 7
 Query Match: 95.80% Indels: 79
 DB: 8 Gaps: 1

US-10-046-935-2234 (1-1116) x ADL91540 (1-450)

QY 1 ATGGACGCTCCGCGTCCGAGAAAGATCTCAGAGTAAAGAACTTAAAGAAATTC 60
 Db 1 MetAspAlaArgValProGlnLysAspLeuArgValLysAsnLeuLysPhe 20
 QY 61 AGATATGTGAAGTGAATTCATGGAACCTCGTCATCTCTGATGACAGTGTGACAGC 120
 Db 21 ArgTyrValLysLeuLysMetGluThrSerSerSerSerSerSerSerSerSerSer 40
 QY 121 TTGCTTCTGATAAATTTTGCAACACG 147
 Db 41 PheAlaSerAspAsnPheAlaAsnThrLysProLysPheArgSerAspLysSerGluGlu 60
 QY 147 147
 Db 61 LeuAlaSerValPheTyrGluAspSerAspAsnGluSerPheCysGlyPheSerGluSer 80
 QY 147 147
 Db 81 GluValGlnAspValLeuAspHisCysGlyPheLeuGlnLysProArgProAspValThr 100
 QY 147 147
 Db 101 AsnGluLeuAlaGlyPheHisAlaAspSerAspAspGluSerPheCysGlyPheSer 120
 QY 148 148
 Db 121 GluSerGluLeuGlnAspGlyMetArgLeuGlnSerValArgGluGlyCysArgThrArg 140
 QY 184 AGCAGTGCAGGCACTCTCGACCTCTCAGGGTGGCGATGAAGTTTCCAGCGGAGTACC 243
 Db 141 SerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSerThr 160
 QY 244 AGGGAGCAACCAAAAAAGCAGAGTCCCGCAGCCCTCAGAGATTCGTGACTGAT 303
 Db 161 ArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThrAsp 180
 QY 304 TCCAACTCCGATTCAGAAGTGAAGTGAATGAATTTTGGAGAAAGGGCTTTAAAT 363
 Db 181 SerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeuAsn 200
 QY 364 ATAAAGCAAAACAAAGCAATGCTTGCAAACTCATGTCTGAATAGAAAGCTTCCCTGGC 423
 Db 201 IleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPheProGly 220
 QY 424 TCGTTCGCGTGAAGACATCCCTCCAGGCTCCGACTCACAATCAAGGAGACCGCAAGG 483
 Db 221 SerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArgArg 240
 QY 484 CGTACATTCGCGGTGTGCTCCAGGAGAAACCTCGAACGAGAGCTCGTCTCTTACC 543
 Db 241 ArgThrPheProGlyValAlaSerArgArgAsnProGluArgAlaArgProLeuThr 260
 QY 544 AGTCAAGTCCCGGATCTCCGGTCCCTGAGCTCTACCCATGGRGAGAGCAGTTTK 603
 Db 261 ArgSerArgSerArgIleLeuLysSerLeuAspAlaLeuProMetGluGluGluGlu 280
 QY 604 GARKMTWGTACATGTTGGTGAAGAGAGAGACCGTGGATGCTACATGAATGAAGAT 663
 Db 281 GluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGluAsp 300
 QY 664 GACCTGCCCAAGACCCGCTCGCTCAGATCATCCGTGACCCCTTCGCATATAATTCGCCCA 723
 Db 301 AspLeuProArgSerArgSerArgSerArgSerValThrLeuProHisIleLeuArgPro 320
 QY 724 GTGGAAGAAATTCAGAGGAGGAGTTCGAGACGCTCGCAATTCCTCGAGAGAGATA 783
 Db 321 ValGluGluIleThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLysIle 340
 QY 784 TATAACCGTTCACTGGGCTCTACTTGTTCATCAATGCCGTGAGAGAGATTTGATACCAA 843
 Db 341 TyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThrLys 360

QY 844 ACAAACTCGAAGAACCCAGACTGCTGGGGCGTTCAGGCCAGTTCCTGTGGCCCTGCCCTT 903
 Db 361 ThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCysLeu 380
 QY 904 CGAAACCGTTATGGTGAAGAGTCAAGGATGCTCTGTGGATCCGAACCTGGCATTCGCCG 963
 Db 381 ArgAsnArgTyrGlyGluGluValArgAspAlaLeuLeuAspProAsnTrpHisCysPro 400
 QY 964 CTTTCTCGAGGAATCTGCAACTGCAAGTTTCTGCCGCGCAGAGATGGACGGTGTGCGACT 1023
 Db 401 ProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAlaThr 420
 QY 1024 GGGTCTCTGTGATTATAGCCAAATATCATGCTTTGGGAATGTGCATCCCTACTTGAA 1083
 Db 421 GlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeuLys 440
 QY 1084 AGCCTGAACACAGGAATTTGAAATGCAAGCA 1113
 Db 441 SerLeuLysGlnGluPheGluMetGlnAla 450
 RESULT 15
 ADN06070
 ID ADN06070 standard; protein; 450 AA.
 XX AC ADN06070;
 XX DT 01-JUL-2004 (first entry)
 XX DE Antipsoriatic protein sequence #1189.
 XX KW antipsoriatic; gene therapy; psoriasis; diagnosis.
 XX OS Homo sapiens.
 XX PN WO2004028479-A2.
 XX PD 08-APR-2004.
 XX PF 25-SEP-2003; 2003WO-US030907.
 XX PR 25-SEP-2002; 2002US-0414006P.
 XX PA (GETH) GENENTECH INC.
 XX PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WJ;
 XX PI Wu TD;
 XX DR WPI; 2004-305105/28.
 XX DR N-PSDB; ADN06069.
 XX PS New PRO nucleic acid or polypeptide, useful for preparing a
 PT pharmaceutical composition for diagnosing or treating psoriasis in a
 PT mammal.
 XX PS Claim 9; SEQ ID NO 2465; 3069pp; English.
 XX CC The invention relates to novel polynucleotide and polypeptides for
 CC treating psoriasis or a sequence having at least 80% identity to the
 CC above sequences. The nucleic acid is useful for preparing a composition
 CC for diagnosing or treating psoriasis in a mammal. This sequence
 CC corresponds to one of the polypeptides of the invention.
 XX SQ Sequence 450 AA;
 Alignment Scores:
 Pred. No.: 3,49e-189 Length: 450
 Score: 1879.50 Matches: 364
 Percent Similarity: 80.89% Conservative: 0
 Best Local Similarity: 80.89% Mismatches: 7
 Query Match: 95.80% Indels: 79
 DB: 8 Gaps: 1
 US-10-046-935-2234 (1-1116) x ADN06070 (1-450)

QY 1 ATGACGCTCCGCGTCCGAGAAAGATCTCAGAGTAAAGAACTTAAAGAAATTC 60
 Db 1 MetAspAlaArgValProGlnLysAspLeuArgValLysLysAsnLeuLysPhe 20
 QY 61 AGATATGTGAAGTTCATTCATCGAAACCTCGTCATCCTCTGATGACAGTTGTGACAGC 120
 Db 21 ArgTyrValLysLeuLysSerMetGluThrSerSerSerSerSerSerSerSerSerSer 40
 QY 121 TTTCTCTCTGATAATTTTGCACACG----- 147
 Db 41 PheAlaSerAspAsnPheAlaAsnThrLysProLysPheArgSerAspLysSerGluGlu 60
 QY 147 ----- 147
 Db 61 LeuAlaSerValPheTyrGluAspSerAspAsnGluSerPheCysGlyPheSerGluSer 80
 QY 147 ----- 147
 Db 81 GluValGlnAspValLeuAspHisCysGlyPheLeuGlnLysProArgProAspValThr 100
 QY 147 ----- 147
 Db 101 AsnGluLeuAlaGlyIlePheHisAlaAspSerAspAspGluSerPheCysGlyPheSer 120
 QY 148 -----AGCTGCAGTCAGTTCGGGAAGGCTGTAGGACCCGC 183
 Db 121 GluSerGluIleGlnAspGlyMetArgLeuGlnSerValArgGluGlyCysArgThrArg 140
 QY 184 AGCCAGTGCAGGCACCTCTGCACCTCTCAGGTCGCGATGAAGTTTCAGCGCGAGTACC 243
 Db 141 SerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSerThr 160
 QY 244 AGGGAGCAACCAACAAAGAGAGTCCCGCAGCCCTCAGAGATTCGTGACCTGAT 303
 Db 161 ArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThrAsp 180
 QY 304 TCCAATCCGATTCAGAAGATGAAGTGAATGAATTTTGGAGAAAGGGCTTTAAAT 363
 Db 181 SerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeuAsn 200
 QY 364 ATAAAGCAACAAACCAATGCTTGCAAACTCATGTCTGAATAGAACTTCCTGCGC 423
 Db 201 IleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPheProGly 220
 QY 424 TCGTTCCTGGAGACATCCCTCCAGGCTCCGACTCACAATCAAGAGACCGCGAAG 483
 Db 221 SerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArgArg 240
 QY 484 CGTACATTCCTGGGTGTTCCTCCAGAGAAACCTCGAACCGAGAGCTCGTCTCTTACC 543
 Db 241 ArgThrPheProGlyValAlaSerArgArgAsnProGluArgAlaArgProLeuThr 260
 QY 544 AGGTCAAGTCCCGGATCCCTCGGTCCTGACGCTCTACCCATGCGAGAGCCAGTTTK 603
 Db 261 ArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGluGlu 280
 QY 604 GARKMTWGTACATGTGTGAGAAAGAGACCGTGGATGGCTACATGAATGAAGAT 663
 Db 281 GluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGluAsp 300
 QY 664 GACCTGCCAGAACCGTCCGATCATCCGTGACCCCTTCGCGATATAATTCGCCCA 723
 Db 301 AspLeuProArgSerArgSerArgSerValThrLeuProHisIleIleArgPro 320
 QY 724 GTGGAGAAATACAGAGGAGGTGGAGACCTCTGACGCAATTCCTCGAGAGAGATA 783
 Db 321 ValGluGluIleThrGluGluLeuGluAsnValCysSerAsnSerArgGluLysIle 340
 QY 784 TATAACCGTTCACCTGGCTCTACTCTCATCAATGCGTCAGAGACTATTGTATCCAAA 843
 Db 341 TyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThrLys 360

QY 844 ARAACTGCAGAAACCCAGACTGCTGGGCGCTTCGAGCCAGTTCTGTGGCCCTGCCTT 903
 Db 361 ThrAsnCysArgAsnProAspCysTyrPglyValArgGlyGlnPheCysGlyProCysLeu 380
 QY 904 CGAAACCGTTATGTGTAAGAGGTCAAGGATGCTCTGCTGGATCCGAACTGCATTGCCCG 963
 Db 381 ArgAsnArgTyrGlyGluValArgAspAlaLeuLeuAspProAsnTrpHisCysPro 400
 QY 964 CTTYGTGAGGAATCTGAACTGCACTGCTGCGGAGCGAGATGACGGTGTGCGACT 1023
 Db 401 ProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAlaThr 420
 QY 1024 GGGTCTCTGTGTATTTAGCCAAATATCATGGCTTTGGGAATGTGCATGCCCTACTTGA 1083
 Db 421 GlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeuLys 440
 QY 1084 AGCCTGAAACAGGAATTTGAAATGCAAGCA 1113
 Db 441 SerLeuLysGlnGluPheGluMetGlnAla 450

Search completed: October 12, 2004, 15:12:46
 Job time : 127 secs

Blank

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame plus n2p model

Run on: October 12, 2004, 15:18:46 ; Search time 100 Seconds
(without alignments)

Title: US-10-046-935-2234

Perfect score: 1962
Sequence: 1 atggacgtctgcgcgcgtgcc.....aattgaaatgcagcataa 1116

Scoring table: BLOSUM62

Xgapop	10.0	Xgapext	0.5
Ygapop	10.0	Ygapext	0.5
Fgapop	6.0	Fgapext	7.0
Delop	6.0	Delext	7.0

Searched: 1356558 seqs, 322682953 residues

Model	Number of hits satisfying chosen parameters	2713116
Model 1	1	1
Model 2	1	1
Model 3	1	1
Model 4	1	1
Model 5	1	1
Model 6	1	1
Model 7	1	1
Model 8	1	1
Model 9	1	1
Model 10	1	1
Model 11	1	1
Model 12	1	1
Model 13	1	1
Model 14	1	1
Model 15	1	1
Model 16	1	1
Model 17	1	1
Model 18	1	1
Model 19	1	1
Model 20	1	1
Model 21	1	1
Model 22	1	1
Model 23	1	1
Model 24	1	1
Model 25	1	1
Model 26	1	1
Model 27	1	1
Model 28	1	1
Model 29	1	1
Model 30	1	1
Model 31	1	1
Model 32	1	1
Model 33	1	1
Model 34	1	1
Model 35	1	1
Model 36	1	1
Model 37	1	1
Model 38	1	1
Model 39	1	1
Model 40	1	1
Model 41	1	1
Model 42	1	1
Model 43	1	1
Model 44	1	1
Model 45	1	1
Model 46	1	1
Model 47	1	1
Model 48	1	1
Model 49	1	1
Model 50	1	1
Model 51	1	1
Model 52	1	1
Model 53	1	1
Model 54	1	1
Model 55	1	1
Model 56	1	1
Model 57	1	1
Model 58	1	1
Model 59	1	1
Model 60	1	1
Model 61	1	1
Model 62	1	1
Model 63	1	1
Model 64	1	1
Model 65	1	1
Model 66	1	1
Model 67	1	1
Model 68	1	1
Model 69	1	1
Model 70	1	1
Model 71	1	1
Model 72	1	1
Model 73	1	1
Model 74	1	1
Model 75	1	1
Model 76	1	1
Model 77	1	1
Model 78	1	1
Model 79	1	1
Model 80	1	1
Model 81	1	1
Model 82	1	1
Model 83	1	1
Model 84	1	1
Model 85	1	1
Model 86	1	1
Model 87	1	1
Model 88	1	1
Model 89	1	1
Model 90	1	1
Model 91	1	1
Model 92	1	1
Model 93	1	1
Model 94	1	1
Model 95	1	1
Model 96	1	1
Model 97	1	1
Model 98	1	1
Model 99	1	1
Model 100	1	1

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

Maximum MACCH 100%
Listing first 45 summaries

Command line parameters:

```

-Model=frame; n2p model -DEV-xlp
-Os/can2_1/USPro_pool_P/US10046935/runat_12102004_155228_15337/app_query.fasta_1.1387
-DB=Published Applications AA -QPM=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blasmus6
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTWTF=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10046935 @CGN_1_153 @runat_12102004_155228_15337
-NCPU=6 -ICPU=3 -NO MWAP -LARGEOQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LOGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

Database : Published Applications AA: *

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/2/pubpaa/US05A_PUBCOMB.pcp.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pcp.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
18: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
19: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB	ID	Description
1	100%	100	100	NCBI	1	...

```

RESULT 1
US-09-878-178-2235
; Sequence 2235, Application US/09878178
; Patent No. US20020177552A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527
; CURRENT APPLICATION NUMBER: US/09/878,178
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 2237
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2235
; LENGTH: 371
; TYPE: prt
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(371)
; OTHER INFORMATION: Xaa = Any amino acid

```

US-09-878-178-2235									
Alignment Scores:									
Pred. No.:	3,06e-178	Length:	371						
Score:	1943.00	Matches:	371						
Percent Similarity:	100.00%	Conservative:	0						
Best Local Similarity:	100.00%	Mismatches:	0						
Query Match:	99.03%	Indels:	0						
DB:	9	Gaps:	0						
US-10-046-935-2234 (1-1116) x US-09-878-178-2235 (1-371)									
QY	1	ATGAGCGCTCGCGCTGCGCGAGAGATCTCAGAGTAAAGAACTTAAAGAAATTC	60						
Db	1	MetAspAlaargValProGlnLysAspLeuValLysAsnLeuLysPhe	20						
QY	61	AGATATGTGAAGTTGATTTCCATGGAACCTCTCATCTCTGATGACAGTGTGACAGC	120						
Db	21	ArgTyrValLysLeuLeuSerMetGluThrSerSerSerAspAspSerCysAspSer	40						
QY	121	TTTGCTTCTGATTAATTTTGCACACAGAGCTGAGTCAGTTCGGGAGGCTGTAGGACC	180						
Db	41	PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr	60						
QY	181	CGCAGCAGTGCAGGCACTCTGGACCTCTCAGGTGCGATGAAGTTTCCAGCGCGGAGT	240						
Db	61	ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer	80						
QY	241	ACCAGGGGAGCAACCAAAAAAGCAGAGTCCCGCAGCCCTCAGAGAAATTCGTGACT	300						
Db	81	ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr	100						
QY	301	GATTCCAACTCCGATTCAGAGATGAAGTGAATGAATTTTGGAGAAAGGGCTTTA	360						
Db	101	AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu	120						
QY	361	AATATAAGCAAAACAAAGCAATCTTGCAAACTCATGTCTGAATPAGAAAGCTCCCT	420						
Db	121	AsnLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro	140						
QY	421	GGCTCGTTCCGTGGAAGACATCCCTCCAGGCTCCGACTCCACATCAAGGAGACCGGA	480						
Db	141	GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg	160						
QY	481	AGCGGTACATTCGCGGTGTTGCTTCCAGAGAAACCTGAAACCGGAGAGCTCGTCTCTT	540						
Db	161	ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgArgAlaArgProLeu	180						
QY	541	ACCAGGTCAAGTCCGAGTCTCGGTCCCTTGCAGCTCTACCCATGGRGAGAGCCAGT	600						
Db	181	ThrArgSerArgSerArgGlyLeuGlySerLeuAspAlaLeuProMet**ArgAlaSer	200						
QY	601	TTKGARKMTWGTACATGTTGGTGAAGAGAGACCGTGGATGGCTACATGAATGAA	660						
Db	201	**Glu***TyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu	220						
QY	661	GATGACCTCCAGAGACCGTCCGTCAGATCATCCGTGACCTTCCGATATAATTCGC	720						
Db	221	AspAspLeuProArgSerArgSerArgSerValThrLeuProHisLleleArg	240						
QY	721	CCAGTGGAGAAATTCAGAGAGAGAGTGGAGAACGCTCTCAGCAATTCCTCGAGAGAG	780						
Db	241	ProValGluGluLleThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLys	260						
QY	781	ATATATAACGTTCACTGGGCTCTACTGTCTCATCAATGCGCTCAGAGAACTATTGATAC	840						
Db	261	IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr	280						
QY	841	AAAAAATCTGCAAAACCCAGACTCTCTGGGCGTTCGAGGCGCAGTTCGTGGCCCTTCG	900						
Db	281	LysThrAsnCysArgAsnProAspCysIleArgValArgGlyGlnPheCysGlyProCys	300						
QY	901	CTTCGAAACCGTTATGGTGAAGAGGTTCAGGATGCTCTCTGATCCGAACTGGCATTCG	960						

US-09-878-178-2235									
Alignment Scores:									
Pred. No.:	3,06e-178	Length:	371						
Score:	1943.00	Matches:	371						
Percent Similarity:	100.00%	Conservative:	0						
Best Local Similarity:	100.00%	Mismatches:	0						
Query Match:	99.03%	Indels:	0						
DB:	9	Gaps:	0						
US-10-046-935-2234 (1-1116) x US-09-878-178-2235 (1-371)									
QY	1	ATGAGCGCTCGCGCTGCGCGAGAGATCTCAGAGTAAAGAACTTAAAGAAATTC	60						
Db	1	MetAspAlaargValProGlnLysAspLeuValLysAsnLeuLysPhe	20						
QY	61	AGATATGTGAAGTTGATTTCCATGGAACCTCTCATCTCTGATGACAGTGTGACAGC	120						
Db	21	ArgTyrValLysLeuLeuSerMetGluThrSerSerSerAspAspSerCysAspSer	40						
QY	121	TTTGCTTCTGATTAATTTTGCACACAGAGCTGAGTCAGTTCGGGAGGCTGTAGGACC	180						
Db	41	PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr	60						
QY	181	CGCAGCAGTGCAGGCACTCTGGACCTCTCAGGTGCGATGAAGTTTCCAGCGCGGAGT	240						
Db	61	ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer	80						
QY	241	ACCAGGGGAGCAACCAAAAAAGCAGAGTCCCGCAGCCCTCAGAGAAATTCGTGACT	300						
Db	81	ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr	100						
QY	301	GATTCCAACTCCGATTCAGAGATGAAGTGAATGAATTTTGGAGAAAGGGCTTTA	360						
Db	101	AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu	120						
QY	361	AATATAAGCAAAACAAAGCAATCTTGCAAACTCATGTCTGAATPAGAAAGCTCCCT	420						
Db	121	AsnLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro	140						
QY	421	GGCTCGTTCCGTGGAAGACATCCCTCCAGGCTCCGACTCCACATCAAGGAGACCGGA	480						
Db	141	GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg	160						
QY	481	AGCGGTACATTCGCGGTGTTGCTTCCAGAGAAACCTGAAACCGGAGAGCTCGTCTCTT	540						
Db	161	ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgArgAlaArgProLeu	180						
QY	541	ACCAGGTCAAGTCCGAGTCTCGGTCCCTTGCAGCTCTACCCATGGRGAGAGCCAGT	600						
Db	181	ThrArgSerArgSerArgGlyLeuGlySerLeuAspAlaLeuProMet**ArgAlaSer	200						
QY	601	TTKGARKMTWGTACATGTTGGTGAAGAGAGACCGTGGATGGCTACATGAATGAA	660						
Db	201	**Glu***TyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu	220						
QY	661	GATGACCTCCAGAGACCGTCCGTCAGATCATCCGTGACCTTCCGATATAATTCGC	720						
Db	221	AspAspLeuProArgSerArgSerArgSerValThrLeuProHisLleleArg	240						
QY	721	CCAGTGGAGAAATTCAGAGAGAGAGTGGAGAACGCTCTCAGCAATTCCTCGAGAGAG	780						
Db	241	ProValGluGluLleThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLys	260						
QY	781	ATATATAACGTTCACTGGGCTCTACTGTCTCATCAATGCGCTCAGAGAACTATTGATAC	840						
Db	261	IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr	280						
QY	841	AAAAAATCTGCAAAACCCAGACTCTCTGGGCGTTCGAGGCGCAGTTCGTGGCCCTTCG	900						
Db	281	LysThrAsnCysArgAsnProAspCysIleArgValArgGlyGlnPheCysGlyProCys	300						
QY	901	CTTCGAAACCGTTATGGTGAAGAGGTTCAGGATGCTCTCTGATCCGAACTGGCATTCG	960						

US-09-878-178-2235									
Alignment Scores:									
Pred. No.:	3,06e-178	Length:	371						
Score:	1943.00	Matches:	371						
Percent Similarity:	100.00%	Conservative:	0						
Best Local Similarity:	100.00%	Mismatches:	0						
Query Match:	99.03%	Indels:	0						
DB:	9	Gaps:	0						
US-10-046-935-2234 (1-1116) x US-09-878-178-2235 (1-371)									
QY	1	ATGAGCGCTCGCGCTGCGCGAGAGATCTCAGAGTAAAGAACTTAAAGAAATTC	60						
Db	1	MetAspAlaargValProGlnLysAspLeuValLysAsnLeuLysPhe	20						
QY	61	AGATATGTGAAGTTGATTTCCATGGAACCTCTCATCTCTGATGACAGTGTGACAGC	120						
Db	21	ArgTyrValLysLeuLeuSerMetGluThrSerSerSerAspAspSerCysAspSer	40						
QY	121	TTTGCTTCTGATTAATTTTGCACACAGAGCTGAGTCAGTTCGGGAGGCTGTAGGACC	180						
Db	41	PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr	60						
QY	181	CGCAGCAGTGCAGGCACTCTGGACCTCTCAGGTGCGATGAAGTTTCCAGCGCGGAGT	240						
Db	61	ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer	80						
QY	241	ACCAGGGGAGCAACCAAAAAAGCAGAGTCCCGCAGCCCTCAGAGAAATTCGTGACT	300						
Db	81	ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr	100						
QY	301	GATTCCAACTCCGATTCAGAGATGAAGTGAATGAATTTTGGAGAAAGGGCTTTA	360						
Db	101	AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu	120						
QY	361	AATATAAGCAAAACAAAGCAATCTTGCAAACTCATGTCTGAATPAGAAAGCTCCCT	420						
Db	121	AsnLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro	140						
QY	421	GGCTCGTTCCGTGGAAGACATCCCTCCAGGCTCCGACTCCACATCAAGGAGACCGGA	480						
Db	141	GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg	160						
QY	481	AGCGGTACATTCGCGGTGTTGCTTCCAGAGAAACCTGAAACCGGAGAGCTCGTCTCTT	540						
Db	161	ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgArgAlaArgProLeu	180						
QY	541	ACCAGGTCAAGTCCGAGTCTCGGTCCCTTGCAGCTCTACCCATGGRGAGAGCCAGT	600						
Db	181	ThrArgSerArgSerArgGlyLeuGlySerLeuAspAlaLeuProMet**ArgAlaSer	200						
QY	601	TTKGARKMTWGTACATGTTGGTGAAGAGAGACCGTGGATGGCTACATGAATGAA	660						
Db	201	**Glu***TyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu	220						
QY	661	GATGACCTCCAGAGACCGTCCGTCAGATCATCCGTGACCTTCCGATATAATTCGC	720						
Db	221	AspAspLeuProArgSerArgSerArgSerValThrLeuProHisLleleArg	240						
QY	721	CCAGTGGAGAAATTCAGAGAGAGAGTGGAGAACGCTCTCAGCAATTCCTCGAGAGAG	780						
Db	241	ProValGluGluLleThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLys	260						
QY	781	ATATATAACGTTCACTGGGCTCTACTGTCTCATCAATGCGCTCAGAGAACTATTGATAC	840						
Db	261	IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr	280						
QY	841	AAAAAATCTGCAAAACCCAGACTCTCTGGGCGTTCGAGGCGCAGTTCGTGGCCCTTCG	900						
Db	281	LysThrAsnCysArgAsnProAspCysIleArgValArgGlyGlnPheCysGlyProCys	300						
QY	901	CTTCGAAACCGTTATGGTGAAGAGGTTCAGGATGCTCTCTGATCCGAACTGGCATTCG	960						

US-09-878-178-2235									
Alignment Scores:									
Pred. No.:	3,06e-178	Length:	371						
Score:	1943.00	Matches:	371						
Percent Similarity:	100.00%	Conservative:	0						
Best Local Similarity:	100.00%	Mismatches:	0						
Query Match:	99.03%	Indels:	0						
DB:	9	Gaps:	0						
US-10-046-935-2234 (1-1116) x US-09-878-178-2235 (1-371)									
QY	1	ATGAGCGCTCGCGCTGCGCGAGAGATCTCAGAGTAAAGAACTTAAAGAAATTC	60						
Db	1	MetAspAlaargValProGlnLysAspLeuValLysAsnLeuLysPhe	20						
QY	61	AGATATGTGAAGTTGATTTCCATGGAACCTCTCATCTCTGATGACAGTGTGACAGC	120						
Db	21	ArgTyrValLysLeuLeuSerMetGluThrSerSerSerAspAspSerCysAspSer	40						
QY	121	TTTGCTTCTGATTAATTTTGCACACAGAGCTGAGTCAGTTCGGGAGGCTGTAGGACC	180						
Db	41	PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr	60						
QY	181	CGCAGCAGTGCAGGCACTCTGGACCTCTCAGGTGCGATGAAGTTTCCAGCGCGGAGT	240						
Db	61	ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer	80						
QY	241	ACCAGGGGAGCAACCAAAAAAGCAGAGTCCCGCAGCCCTCAGAGAAATTCGTGACT	300						
Db	81	ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr	100						
QY	301	GATTCCAACTCCGATTCAGAGATGAAGTGAATGAATTTTGGAGAAAGGGCTTTA	360						
Db	101	AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu	120						
QY	361	AATATAAGCAAAACAAAGCAATCTTGCAAACTCATGTCTGAATPAGAAAGCTCCCT	420						
Db	121	AsnLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro	140						
QY	421	GGCTCGTTCCGTGGAAGACATCCCTCCAGGCTCCGACTCCACATCAAGGAGACCGGA	480						
Db	141	GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg	160						
QY	481	AGCGGTACATTCGCGGTGTTGCTTCCAGAGAAACCTGAAACCGGAGAGCTCGTCTCTT	540						
Db	161	ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgArgAlaArgProLeu	180						
QY	541	ACCAGGTCAAGTCCGAGTCTCGGTCCCTTGCAGCTCTACCCATGGRGAGAGCCAGT	600						
Db	181	ThrArgSerArgSerArgGlyLeuGlySerLeuAspAlaLeuProMet**ArgAlaSer	200						
QY	601	TTKGARKMTWGTACATGTTGGTGAAGAGAGACCGTGGATGGCTACATGAATGAA	660						
Db	201	**Glu***TyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu	220						
QY	661	GATGACCTCCAGAGACCGTCCGTCAGATCATCCGTGACCTTCCGATATAATTCGC	720						
Db	221	AspAspLeuProArgSerArgSerArgSerValThrLeuProHisLleleArg	240						
QY	721	CCAGTGGAGAAATTCAGAGAGAGAGTGGAGAACGCTCTCAGCAATTCCTCGAGAGAG	780						
Db	241	ProValGluGluLleThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLys	260						
QY	781	ATATATAACGTTCACTGGGCTCTACTGTCTCATCAATGCGCTCAGAGAACTATTGATAC	840						
Db	261	IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr	280						
QY	841	AAAAAATCTGCAAAACCCAGACTCTCTGGGCGTTCGAGGCGCAGTTCGTGGCCCTTCG	900						
Db	281	LysThrAsnCysArgAsnProAspCysIleArgValArgGlyGlnPheCysGlyProCys	300						
QY	901	CTTCGAAACCGTTATGGTGAAGAGGTTCAGGATGCTCTCTGATCCGAACTGGCATTCG	960						

US-09-878-178-2235									
Alignment Scores:									
Pred. No.:	3,06e-178	Length:	371						
Score:	1943.00	Matches:	371						
Percent Similarity:	100.00%	Conservative:	0						
Best Local Similarity:	100.00%	Mismatches:	0						
Query Match:	99.03%	Indels:	0						
DB:	9	Gaps:	0						
US-10-046-935-2234 (1-1116) x US-09-878-178-2235 (1-371)									
QY	1	ATGAGCGCTCGCGCTGCGCGAGAGATCTCAGAGTAAAGAACTTAAAGAAATTC	60						
Db	1	MetAspAlaargValProGlnLysAspLeuValLysAsnLeuLysPhe	20						
QY	61	AGATATGTGAAGTTGATTTCCATGGAACCTCTCATCTCTGATGACAGTGTGACAGC	120						
Db	21	ArgTyrValLysLeuLeuSerMetGluThrSerSerSerAspAspSerCysAspSer	40						
QY	121	TTTGCTTCTGATTAATTTTGCACACAGAGCTGAGTCAGTTCGGGAGGCTGTAGGACC	180						
Db	41	PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr	60						
QY	181	CGCAGCAGTGCAGGCACTCTGGACCTCTCAGGTGCGATGAAGTTTCCAGCGCGGAGT	240						
Db	61	ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer	80						
QY	241	ACCAGGGGAGCAACCAAAAAAGCAGAGTCCCGCAGCCCTCAGAGAAATTCGTGACT	300						
Db	81	ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr	100						
QY	301	GATTCCAACTCCGATTCAGAGATGAAGTGAATGAATTTTGGAGAAAGGGCTTTA	360						
Db	101	AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu	120						
QY	361	AATATAAGCAAAACAAAGCAATCTTGCAAACTCATGTCTGAATPAGAAAGCTCCCT	420						
Db	121	AsnLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro	140						
QY	421	GGCTCGTTCCGTGGAAGACATCCCTCCAGGCTCCGACTCCACATCAAGGAGACCGGA	480						
Db	141	GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg	160						
QY	481	AGCGGTACATTCGCGGTGTTGCTTCCAGAGAAACCTGAAACCGGAGAGCTCGTCTCTT	540						
Db	161	ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgArgAlaArgProLeu	180						
QY	541	ACCAGGTCAAGTCCGAGTCTCGGTCCCTTGCAGCTCTACCCATGGRGAGAGCCAGT	600						
Db	181	ThrArgSerArgSerArgGlyLeuGlySerLeuAspAlaLeuProMet**ArgAlaSer	200						
QY	601	TTKGARKMTWGTACATGTTGGTGAAGAGAGACCGTGGATGGCTACATGAATGAA	660						
Db	201	**Glu***TyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu	220						
QY	661	GATGACCTCCAGAGACCGTCCGTCAGATCATCCGTGACCTTCCGATATAATTCGC	720						
Db	221	AspAspLeuProArgSerArgSerArgSerValThrLeuProHisLleleArg	240						
QY	721	CCAGTGGAGAAATTCAGAGAGAGAGTGGAGAACGCTCTCAGCAATTCCTCGAGAGAG	780						
Db	241	ProValGluGluLleThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLys	260						
QY	781	ATATATAACGTTCACTGGGCTCTACTGTCTCATCAATGCGCTCAGAGAACTATTGATAC	840						
Db	261	IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr	280						
QY	841	AAAAAATCTGCAAAACCCAGACTCTCTGGGCGTTCGAGGCGCAGTTCGTGGCCCTTCG	900						
Db	281	LysThrAsnCysArgAsnProAspCysIleArgValArgGlyGlnPheCysGlyProCys	300						
QY	901	CTTCGAAACCGTTATGGTGAAGAGGTTCAGGATGCTCTCTGATCCGAACTGGCATTCG	960						

US-09-878-178-2235									
Alignment Scores:									
Pred. No									

RESULT 2

```

US-10-046-935-2235
; Sequence 2235, Application US/10046935
; Publication No. US2002015601A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527C1
; CURRENT APPLICATION NUMBER: US/10/046,935
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2239
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2235
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 197, 201, 203, 204
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-046-935-2235

Alignment Scores:
Pred. No.:      3.06e-178      Length:      371
Score:          1943.00      Matches:     371
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00%  Mismatches:  0
Query Match:     99.03%      Indels:      0
DB:              13          Gaps:        0

US-10-046-935-2234 (1-1116) x US-10-046-935-2235 (1-371)

QY      1  ATGACGCTCGCGCTGCGCGAGAGATCTCAGAGTAAAGAACTTAAAGAAATTC  60
Db      1  MetAspAlaArgValProGlnLysAspLeuArgValLysLysAsnLeuLysLysPhe  20

QY     61  AGATATGTGAAGTTGATTTCCATGGAACCTCTCATCTCTGATGACAGTGTGACAGC  120
Db     21  ArgTyrValLysLeuLleSerMetGluThrSerSerSerAspAspSerCysAspSer  40

QY    121  TTTGCTTCTGATTAATTTTGCACACAGAGCTGCAGTCAGTTCGGGAAGGCTCTAGGACC  180
Db    41  PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr  60

QY    181  CGCAGCCAGTGCAGGCACTCTGGACCTCTCAGGGTGGCGATGAAGTTTCCAGCGCGAGT  240
Db    61  ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer  80

QY    241  ACCAGGGGAGCAACCAAAAAAGCAGAGTCCCGCCAGCCCTCAGAGAAATTCGTGACT  300
Db    81  ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr  100

QY    301  GATTCCAACTCCGATTCAGAGATGAAGTGGAAATGAATTTTTCGAGAAAAGGGCTTTA  360

```



```
Db 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 120
QY 361 AATATAAAGCAAAACAAAGCAATGCTTGCACAACTCATGCTCTGAATTAGAAAGCTCCCT 420
Db 121 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140
QY 421 GGCTGCTCCGTGGGAAGACATCCCTCCAGCTCCGACTCACAATCAAGGAGACCGGA 480
Db 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160
QY 481 AGCGTACATCCCGGTGTGCTTCCAGGAGAAACCTGACGAGAGCTGCTCTT 540
Db 161 ArgArgThrPheProGlyValAlaSerArgAsnProGluArgAlaArgProLeu 180
QY 541 ACCAGGTCAAGTCCCGGATCTCCGGTCCCTTGACGCTCTACCCATGGRGAGCCAGT 600
Db 181 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMet**ArgAlaSer 200
QY 601 TTGARKMTWGTACATGTTGGTGAAGAGAGACCGGTGATGGCTACATGAATGAA 660
Db 201 ***Glu****TyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220
QY 661 GATGACCTGCCAGAGCCGCTGCTCCAGATCATCCGTGACCTTCCGCATATAATTCG 720
Db 221 AspAspLeuProArgSerArgSerArgSerValThrLeuProHisIleIleArg 240
QY 721 CCAGTGAAGAAATACAGAGAGAGAGTTGGAGAACGCTCTCAGCAATCTCGAGAGAG 780
Db 241 ProValGluGluIleThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLys 260
QY 781 ATATATAACCGTTCACTGGGCTCTACTTGTATCAATGCCGTGAGAGACTATTGATACC 840
Db 261 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 280
QY 841 AAAAACAACAGCAAAACCCAGACTGCTGGGCGCTTGAGGCCAGTCTGTGGCCCTTCG 900
Db 281 LysThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCys 300
QY 901 CTTGGAACCGTTATGCTGAAGAGTCAGGATCGTCTGCTGGATCCGAACTGCGATTCG 960
Db 301 LeuArgAsnArgTyrGlyGluGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 320
QY 961 CCGCTTGTGAGAAATCTGCAACTGCACTTCTGCGGAGGAGATGACCGGTGTGCG 1020
Db 321 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 340
QY 1021 ACTGGGCTCTGTGATTTAGCCMAATATCATGCGCTTTGGGAATGCGATGCGCTACTTG 1080
Db 341 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 360
QY 1081 AAAAGCTGAACAGGAATTTGAAATGCAAGCA 1113
Db 361 LysSerLeuLysGlnGluPheGluMetGlnAla 371
```

RESULT 3

```
US-10-146-502-2235
; Sequence 2235, Application US/10146502
; Publication No. US20030069180A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.527C2
; CURRENT APPLICATION NUMBER: US/10/146,502
; CURRENT FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2235
```

```
; LENGTH: 371
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 197, 201, 203, 204
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-146-502-2235

Alignment Scores:
Pred. No.: 3,06e-178 Length: 371
Score: 1943.00 Matches: 371
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.03% Indels: 0
DB: 14 Gaps: 0

US-10-046-935-2234 (1-1116) x US-10-146-502-2235 (1-371)
```

```
QY 1 ATGGACGCTCCCGGTGCCCGCAAGAGATCTCAGAGTAAAGAGAACTTAAAGAAATTC 60
Db 1 MetAspAlaArgArgValProGlnLysAspLeuArgValLysLysAsnLeuLysLysPhe 20
QY 61 AGATATGTGAAGTTGATTTCATCGAAACCTCGTCATCCTCTGATGACAGTTGTGACAGC 120
Db 21 ArgTyrValLysLeuIleSerMetGluThrSerSerSerSerSerSerSerCysAspSer 40
QY 121 TTTCTCTTGTATAATTTTGCACACACGAGGCTGCGAGTCAGTTCCGGAAGGCTGTAGGACC 180
Db 41 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 60
QY 181 CGCAGCCAGTGCAGGCACTCTGGACCTCTCAGGGTGGCGATGAAGTTTCCAGCCGGAGT 240
Db 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80
QY 241 ACCAGGGGAGCAACCAACAAAGAGAGAGTCCCGCCAGCCCTCAGAGAATCTGTGACT 300
Db 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100
QY 301 GATTCCAACTCCGATTCAGAAAGATGAAGTGAATGAATTTTGGAGAAAAGGCTTTA 360
Db 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 120
QY 361 AATATAAAGCAAAACAAAGCAATGCTTGCACAACTCATGCTCTGAATTAGAAAGCTTCCT 420
Db 121 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140
QY 421 GGCTGCTCCGTGGAAGACATCCCTCCAGGCTCCGACTCACAATCAAGGAGACCGGA 480
Db 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160
QY 481 AGCGTACATCCCGGTGTGCTTCCAGGAGAAACCTGACGAGAGCTCGTCTCTT 540
Db 161 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgAlaArgProLeu 180
QY 541 ACCAGGTCAAGTCCCGGATCTCCGGTCCCTTGACGCTCTACCCATGGRGAGCCAGT 600
Db 181 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMet**ArgAlaSer 200
QY 601 TTGARKMTWGTACATGTTGGTGAAGAGAGACCGGTGATGGCTACATGAATGAA 660
Db 201 ***Glu****TyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220
QY 661 GATGACCTGCCAGAGCCGCTGCTCCAGATCATCCGTGACCTTCCGCATATAATTCG 720
Db 221 AspAspLeuProArgSerArgSerArgSerValThrLeuProHisIleIleArg 240
QY 721 CCAGTGAAGAAATACAGAGAGAGAGTTGGAGAACGCTCTCAGCAATCTCGAGAGAG 780
Db 241 ProValGluGluIleThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLys 260
QY 781 ATATATAACCGTTCACTGGGCTCTACTTGTATCAATGCCGTGAGAGACTATTGATACC 840
```

Db 261 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 280
QY 841 AAAACAACTGCAGAAACCCAGACTGCTGGGGCGTTCGAGGCCAGTTCTGTGGCCCTGC 900
Db 281 LysThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCys 300
QY 901 CTTCCGAAACCGTTATGGTCAACAGGTCAGGGATGCTCTCTCGATCCGAACTGGCATTGC 960
Db 301 LeuArgAsnArgTrpGlyGlnGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 320
QY 961 CCGCTTGTCCGAGGAATTCGCAACTGCACTTCTGCCGCGCAGCAGATGAGCGGTGCG 1020
Db 321 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 340
QY 1021 ACTGGGCTCTGTGTATTAGCCAAATATCATGGCTTTGGGAAATGTCATGCTGCTACTTG 1080
Db 341 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 360
QY 1081 AAAAGCCTGAAACAGGAATTTGAAATGCAAGCA 1113
Db 361 LysSerLeuLysGlnGluPheGluMetGlnAla 371

RESULT 4

US-10-066-543-1422

; Sequence 1422, Application US/10066543

; Publication No. US20030087818A1

; GENERAL INFORMATION:

; APPLICANT: Jiang, Yuqiu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margareta
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE OF INVENTION: AND DIAGNOSIS OF COLON CANCER

; FILE REFERENCE: 210121.563

; CURRENT APPLICATION NUMBER: US/10/066,543

; CURRENT FILING DATE: 2002-01-31

; NUMBER OF SEQ ID NOS: 3417

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1422

; LENGTH: 371

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: VARIANT

; LOCATION: 197,201,203,204

; OTHER INFORMATION: Xaa = Any amino acid

US-10-066-543-1422

Alignment Scores:
Pred. No.: 3,06e-178 Length: 371
Score: 1943.00 Matches: 371
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.03% Indels: 0
DB: 14 Gaps: 0

US-10-046-935-2234 (1-1116) x US-10-066-543-1422 (1-371)

QY 1 ATGACCGTCCCGCGTCCGAGAAAGATCTCAGAGTAAGAGAACTTAAGAAATTC 60
Db 1 MetAspAlaArgValProGlnLysAspLeuArgValLysLysAsnLeuLysPhe 20
QY 61 AGATATGTGAAGTTCATTTCCATGGAACCTCGTCATCTCTGATGACAGTGTGACAGC 120
Db 21 ArgTyrValLysLeuIleSerMetGluThrSerSerSerAspSerCysAspSer 40

QY 121 TTTGCTCTCTGATAAATTTTGCACAAACAGAGGCTGCAGTCAGTTCCGGAGCGCTGTAGACC 180
Db 41 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGlnGlyCysArgThr 60
QY 181 CGACGCCAGTGCAGCACACTCTCGACTCTCAGGGTGGCGATGAAGTTTCCAGCGCGAGT 240
Db 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80
QY 241 ACCAGGGAGCAACCAACAAAAAGCAGAGTCCCGCCAGCCCTCAGAGAAATTTCTGTGACT 300
Db 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100
QY 301 GATTCAACTCCGATTCAGAACTGAAGTGAATGAATTTTGGAGAAAAGGGCTTTTA 360
Db 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGlnLysArgAlaLeu 120
QY 361 ATATATAAGCAAAACAAACCAATGCTTCGAAAACATCATGCTGAAATAGAAAAGCTTCCCT 420
Db 121 AsnLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140
QY 421 GGCTCGTTCGGTGAAGACATCCCTCCAGGCTCCGACTCACAATCAAGGAGACCGCGA 480
Db 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160
QY 481 AGCGTACATTCGCGGTGTGCTTCAGGAGAAACCCCTGAACGAGAGAGCTCCTCTT 540
Db 161 ArgArgThrPheProGlyValAlaSerArgAsnProGlnArgAlaArgProLeu 180
QY 541 ACCAGGTCAAGTCCCGGATCCCTCGGTCCCTTGAGGCTCTACCCATGGRGAGAGCCAGT 600
Db 181 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMet***ArgAlaSer 200
QY 601 TTXGARMTMWGTACATGTTGGTGAGAAAGAGAACCGGTGATGCTCATGAATGAA 660
Db 201 ***Glu***TyrMetLeuValArgLysThrValAspGlyTyrMetAsnGlu 220
QY 661 GATGACCTGCCAGAGCCGCTCGCTCCAGATCATCGTGACCTCCCGCATATATAATTCG 720
Db 221 AspAspLeuProArgSerArgArgSerValThrLeuProHisIleLeuArg 240
QY 721 CCAGTGAAGAAATTCAGAGAGAGAGTTCGAGAAACGCTCTGCAGCAATTCGAGAGAAG 780
Db 241 ProValGluGlnIleThrGluGluGluLeuAsnValCysSerAsnSerArgGluLys 260
QY 781 ATATATAACCGTTCACCTGGCTCTACTTGTCAATGATCCGTCAGAGACTATTGATACC 840
Db 261 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 280
QY 841 AAAACAACTGCAGAAACCCAGACTGCTGGGGCGTTCGAGCCAGTTCTGTGGCCCTGC 900
Db 281 LysThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCys 300
QY 901 CTTGAAAACCGTTATGTGTGAAGAGTTCAGGGATGCTCTGTGATCCGAACTGCAGTTCG 960
Db 301 LeuArgAsnArgTyrGlyGluGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 320
QY 961 CCGCTTGTTCAGGAAATCTGCAACTGAGTTTCTCGCGGAGCAGAGTGGAGCGTGTGCG 1020
Db 321 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 340
QY 1021 ACTGGGCTCTGTGTATTAGCCAAATATCATGCTTTGGAAATGTGCATGCGCTTACTTG 1080
Db 341 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 360
QY 1081 AAAAGCCTGAAACAGGAATTTGAAATGCAAGCA 1113
Db 361 LysSerLeuLysGlnGluPheGluMetGlnAla 371

RESULT 5

US-10-301-822-16

; Sequence 16, Application US/10301822

; Publication No. US20030148410A1

; GENERAL INFORMATION:

```
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MM001-029P2RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,989
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-10-301-822-16

Alignment Scores:
Pred. No.: 6.87e-177 Length: 371
Score: 1929.00 Matches: 364
Percent Similarity: 98.11% Conservative: 0
Best Local Similarity: 98.11% Mismatches: 7
Query Match: 98.32% Indels: 0
DB: 14 Gaps: 0

US-10-046-935-2234 (1-1116) x US-10-301-822-16 (1-371)

QY 1 ATGACGCTCCGCGCTGCGCGAGAAAGATCTCAGAGTAAAGAAAGAACTTAAAGAAATTC 60
Db 1 MetAspAlaArgValProGlnLysAspLeuArgValLysLysAsnLeuLysLysPhe 20

QY 61 AGATATGTGAAGTTGATTTCATGGAAACCTCGTCATCCTCTGATGACAGTTGTGACAGC 120
Db 21 ArgTyrValLysLeuIleSerMetGluThrSerSerSerSerSerSerSerSerSerSer 40

QY 121 TTGTCTCTCATATATTTTGCACACACAGAGTGCAGTCAGTCGGAAGGCTGTAGGACC 180
Db 41 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 60

QY 181 CGCAGCCAGTGCAGGCACTCTGGACCTCTCAGGTTGGCGATGAAGTTTCCAGCGCGGAGT 240
Db 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80

QY 241 ACCAGGGGACCAACCAACAAAGAGAGTCCCGCCAGCCCTCAGAGAAATTTCTGTGACT 300
Db 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100

QY 301 GATTCCAACCTCCGATTCAGAGATGAAGTGAATGAATTTTGGAGAAAGGCTTTA 360
Db 101 AspSerAsnSerSerSerSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 120

QY 361 AATATAAGCAAAACAAAGCAATCTCTGCAAAACTCATGTCTGAATTAGAAAGCTTCCCT 420
Db 121 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140

QY 421 GGCTCTTCGTTGGAAGACATCCCTCCAGGCTCCGACTCAATCAAGAGACCGCGA 480
Db 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgProArg 160

QY 481 AGCGGTACATTCGCGGTGTGTCTCCAGGAGAAACCTGCAAGCGAGAGCTGCTCTT 540
Db 161 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgArgAlaArgProLeu 180
```

```
QY 541 ACCAGGTCAAGTCCCGATCTCTGGGTCCCTTGAGCTCTACCCATGGRGAGAGCCAGT 600
Db 181 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGlu 200

QY 601 TTGARKMTWMTGATCATGTTGGTGAAGAGGAGACCGTGGATGGCTACATGAATGAA 660
Db 201 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220

QY 661 GATGACCTGCCCAAGCCGCTCCAGATCATCCGTGACCCCTTCCGCATATAATTCGC 720
Db 221 AspAspLeuProArgSerArgArgSerValThrLeuProHisIleIleArg 240

QY 721 CCAGTGGAGAAATACAGAGAGAGATTGGAGAACGTCCTGCAGCAATTCGAGAGAG 780
Db 241 ProValGluGluIleThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLys 260

QY 781 ATATATAACCGTTCACCTGGGCTCTACTTGTCTATCAATCAATCCGCTCAGAGACTATTGATACC 840
Db 261 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 280

QY 841 AAAACAAACTGCAGAAACCCAGACTGCTGGGGCGTTTCAGGCCAGTTCCTGTGGCCCTGC 900
Db 281 LysThrAsnCysArgAsnProAspCysTyrpGlyValArgGlyGlnPheCysGlyProCys 300

QY 901 CTTCCGAAACCGTTATGCTGAGAGAGTCAGGATGCTCTGCTGGATCCGAACTGCATTGC 960
Db 301 LeuArgAsnArgTyrGlyGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 320

QY 961 CCGCTTTGTCGAGGAATCTGCACTGCAGTTTCTGCCGCGAGCAGAGATGACGGTGTGCG 1020
Db 321 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 340

QY 1021 ACTGGGTCTCTGTGTTATTTAGCCAAATATCATGGCTTTGGGATGTCATGCTACTTGT 1080
Db 341 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 360

QY 1081 AAAAGCTGAAACAGGAATTTGAAATGCAAGCA 1113
Db 361 LysSerLeuLysGlnGluPheGluMetGlnAla 371

RESULT 6
US-10-295-027-126
; Sequence 126, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
```

```
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 126
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-126

Alignment Scores:
Pred. No.: 6,87e-177      Length: 371
Score: 1929.00           Matches: 364
Percent Similarity: 98.11% Conservative: 0
Best Local Similarity: 98.11% Mismatches: 7
Query Match: 98.32%      Indels: 0
DB: 14                  Gaps: 0

US-10-046-935-2234 (1-1116) x US-10-295-027-126 (1-371)

QY 1 ATGGAGCTGCGCGGTGGCGAAGAAGATCTCAGAGTAAGAAGAACTTAAAGAATTC 60
Db 1 MetAspAlaArgValProGlnLysAspLeuArgValLysLysLeuLysLysPhe 20

QY 61 AGATATGTGAAGTTGATTTCCATGGAACCTCGTCATCTCATGACAGTGTGACAGC 120
Db 21 ArgTyrValLysLeuLeuSerMetGluThrSerSerSerSerSerSerSerSerSer 40

QY 121 TTTCCTCTCATAAATTTTGAACACAGAGCTCGAGTCAGTTCGGGAAGGCTGTAGAAC 180
Db 41 PheAlaSerAspAsnPheAlaAenThrArgLysGlnSerValArgGluGlyCysArgThr 60

QY 181 CGAGCAGCTGCGAGCACTCGACCTCTCAGGTGCGGATGAAGTTTCAGCGCGGAGT 240
Db 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80

QY 241 ACCAGGGAGCAACCAACAAAGAGAGAGTCCCGCCAGCCCTCAGAGAAATTCGTGACT 300
Db 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100

QY 301 GATTCGAATCCGATTCAGAAATGAAGTGGAAATGAATTTTGGAGAAAAGGGCTTTA 360
Db 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 120

QY 361 AATATAAGCAAAACAAGCAATCTTGCAGAACTCATGCTGAATTAGAAGCTTCCT 420
Db 121 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140

QY 421 GGCTCGTTTCGTGGAAGACATCCCTCCAGGCTCCGACTCACAAATCAAGAGACCGGA 480
Db 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160

QY 481 AGCGGTACATTCGCGGGTGTTCCTCCAGGAGAAACCTGAAACCGAGAGCTCGTCTT 540
Db 161 ArgArgThrPheProGlyValAlaSerArgAsnProGluArgArgAlaArgProLeu 180

QY 541 ACCAGGTCAGGTCCTCGGTCCCTTCAGCGCTCTACCATTCGGRGAGAGCCAGT 600
Db 181 ThrArgSerArgSerArgIleuGlySerLeuAspAlaLeuProMetGluGluGlu 200

QY 601 TTKGARKMTWMTACATGTTGTGAGAAAGAGAGACCGGTGGATGGCTACATGAATGAA 660
Db 201 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220

QY 661 GATGACCTGCCAGAGCGGTGCTCCAGATCATCCGTGACCCCTCCGCATATAATTCGC 720
Db 121 TTTCCTCTCATAAATTTTGAACACAGAGCTCGAGTTCGGGAAGGCTGTAGAAC 180
```

```
221 AspAspLeuProArgSerArgSerArgSerValThrLeuProHisIleLeuArg 240
721 CCAGTGGGAAGAAATTACAGAGGAGGAGTTGGAGAACGCTCGACCAATTCGAGAGAAG 780
241 ProValGluGluIleThrGluGluLeuGluAsnValCysSerAsnSerArgGluLys 260
781 ATATATAACCGTTCACTGGGCTCTACTTGTCTCATCAATGCCGTCAGAGACTATTGATACC 840
261 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 280
841 AAAACAAACTGCAGAAACCCAGACAGCTGCTGGGGCGTTTCGAGGCCAGGTTCTGTGCCCTGC 900
281 LysThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCys 300
901 CTTGGAACCGTTATGCTGAAGAGGTGAGGATGCTCTCTGTCGATCCGAACTGGCATTCG 960
301 LeuArgAsnArgTyrGlyGluValAlaArgAspAlaLeuLeuAspProAsnTrpHisCys 320
961 CCGCCTTGTGAGGAATCTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
321 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 340
1021 ACTGGGCTCTGTGTATTTAGCCAAATATCATGGCTTTGGGAATGTGATGCTGCTACTTG 1080
341 ThrGlyValLeuValLysLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 360

1081 AAAAGCCTGAACAGGAATTTGAAATGCAAGCA 1113
361 LysSerLeuLysGlnGluPheGluMetGlnAla 371

RESULT 7
US-10-046-935-2239
; Sequence 2239, Application US/10046935
; Publication No. US20020156011A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527C1
; CURRENT APPLICATION NUMBER: US/10/046,935
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2239
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 2239
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-046-935-2239

Alignment Scores:
Pred. No.: 7,02e-177      Length: 391
Score: 1929.00           Matches: 364
Percent Similarity: 98.11% Conservative: 0
Best Local Similarity: 98.11% Mismatches: 7
Query Match: 98.32%      Indels: 0
DB: 13                  Gaps: 0

US-10-046-935-2234 (1-1116) x US-10-046-935-2239 (1-391)

QY 1 ATGACGCTCGCCCGTGGCGAAGATCTCAGATAAGAAAGAACTTAAAGAAATTC 60
Db 21 MetAspAlaArgArgValProGlnLysAspLeuArgValLysLysLeuLysLysPhe 40

QY 61 AGATATGTGAAGTTGATTTCCATGGAACCTCGTCATCTCATGACAGTGTGACAGC 120
Db 41 ArgTyrValLysLeuLeuSerMetGluThrSerSerSerSerSerSerSerSerSer 60

121 TTTCCTCTCATAAATTTTGAACACAGAGCTCGAGTTCGGGAAGGCTGTAGAAC 180
```

```
Db 61 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 80
QY 181 CGCAGCCAGTCAGGCACTCTGGACCTCTCAGGTTGGCGATGAAGTTTCCAGCGCGGAGT 240
Db 81 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 100
QY 241 ACCAGGGAGGACCAACAAACAAAAGCAGAGTCCGCCAGCCCTCAGAGAATTTCTGTGACT 300
Db 101 ThrArgGlyAlaThrAsnLysAlaGluSerArgGlnProSerGluAsnSerValThr 120
QY 301 GATTCCCACTCCGATTCAGAGATCAAGTGAATTAATTTTGGAGAAAAGGCTTTA 360
Db 121 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 140
QY 361 AATAAAGCAAAACAAAGCAATGCTTCAAACTCATGTCTGAATTAAGAAAGTTCCT 420
Db 141 AsnLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 160
QY 421 GGCTGCTTCGTGGAGACATCCCTCCAGCTCCGACTCACATCAAGGAGACCGGA 480
Db 161 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgProArg 180
QY 481 AGCGGTACATTCGCGGTGTTCTCCAGGAGAAACCTGACGAGAGCTCGTCTCTT 540
Db 181 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgAlaArgProLeu 200
QY 541 ACCAGTCAAGGTCCTCGGATCCTCGGTCCTTCAAGCTCTACCCATGGRGAGACCCAGT 600
Db 201 ThrArgSerArgSerArgGlyLeuGlySerLeuAspAlaLeuProMetGluGluGlu 220
QY 601 TTKGARKMTWMTACATGTTGGTGAGAAAGAGACCGGTGGATGGCTACATGAATCAA 660
Db 221 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 240
QY 661 GATGACCTGCCAGAACCGTCGCTCCAGATCATCCGTGACCCCTCCGCAATATAATTCG 720
Db 241 AspAspLeuProArgSerArgSerArgSerValThrLeuProHisIleArg 260
QY 721 CCAGTGAAGAAATACAGAGGAGGAGTGGAGAACGTCGACAGCAATTCGAGAGAGAG 780
Db 261 ProValGluGluIleThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLys 280
QY 781 ATATATAACCGTTCACTGGGCTTACTTGTCAATCAATGCCGTGAGAGACTATTGATACC 840
Db 281 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 300
QY 841 AAAACAACTGCAGAACCCAGACTGCTGGGGGCTTCAGGCCAGTCTGTGGCCCTGCG 900
Db 301 LysThrAsnCysArgAsnProAspCysTyrPGLyValArgGlyGlnPheCysGlyProCys 320
QY 901 CTTTCAAAACCGTTATGTTGAAGAGGTGAGGATGCTCTGTGGATCGAACTGGCATTGC 960
Db 321 LeuArgAsnArgTyrGlyGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 340
QY 961 CGCCTTGTGAGAGATCTGCAACTGCACTGTTTGGCGGAGGAGATGACCGTGTGG 1020
Db 341 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 360
QY 1021 ACTGGGTCCTCTGTATTTAGCCAATATCATGGCTTTGGATGTGGCATGCTACTTG 1080
Db 361 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 380
QY 1081 AAAAGCCTGAAACAGGAATTTGAAATCAAGCA 1113
Db 381 LysSerLeuLysGlnGluPheGluMetGlnAla 391
```

RESULT 8

```
US-10-146-502-2239
; Sequence 2239, Application US/10146502
; Publication No. US20030069180A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yudi
; APPLICANT: Harlocker, Susan L.
```

```
; APPLICANT: Secretist, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527C2
; CURRENT APPLICATION NUMBER: US/10/146.502
; CURRENT FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2239
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-502-2239

Alignment Scores:
Pred. NO.: 7.02e-177 Length: 391
Score: 1929.00 Matches: 364
Percent Similarity: 98.11% Conservative: 0
Best Local Similarity: 98.11% Mismatches: 7
Query Match: 98.32% Indels: 0
DB: 14 Gaps: 0
US-10-046-935-2234 (1-1116) x US-10-146-502-2239 (1-391)

QY 1 ATGACGCTCGCCGTCGCGAGAAAGATCTCAGAGTAAGAGAACTTAAGAAATTC 60
Db 21 MetAspAlaArgArgValProGlnLysAspLeuArgValLysLysAsnLeuLysLysPhe 40
QY 61 AGATATCTGAAGTTGATTTCCATGGAACCTCGTCATCTCTGATGACAGTTGTGACAGC 120
Db 41 ArgTyrValLysLeuIleSerMetGluThrSerSerSerSerSerSerSerSerSer 60
QY 121 TTTGCTTCTGATAATTTTGCAACACAGAGCTGAGTCCAGTTCGGGAGGCTGTAGGACC 180
Db 61 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 80
QY 181 CGCAGCCAGTCAGGCACTCTGGACCTCTCAGGTTGGCGATGAAGTTTCCAGCGCGGAGT 240
Db 81 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 100
QY 241 ACCAGGGAGGACCAACAAACAAAAGCAGAGTCCGCCAGCCCTCAGAGAATTTCTGTGACT 300
Db 101 ThrArgGlyAlaThrAsnLysAlaGluSerArgGlnProSerGluAsnSerValThr 120
QY 301 GATTCCCACTCCGATTCAGAGATGAAGTGAATTAATTTTGGAGAAAAGGCTTTA 360
Db 121 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 140
QY 361 AATAAAGCAAAACAAAGCAATGCTTGGCAAACTCATGTCTGAATTAAGAAAGTTCCT 420
Db 141 AsnLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 160
QY 421 GGCTGCTTCGTGGAGACATCCCTCCAGGTCCTCGGATGCTTGGAGAAAAGGCTTTA 480
Db 161 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgProArg 180
QY 481 AGCGGTACATTCGCGGTGTTCTCCAGGAGAAACCTGACGAGAGCTCGTCTCTT 540
Db 181 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgAlaArgProLeu 200
QY 541 ACCAGTCAAGGTCCTCGGATCCTCGGTCCTTCAAGCTCTACCCATGGRGAGACCCAGT 600
Db 201 ThrArgSerArgSerArgGlyLeuGlySerLeuAspAlaLeuProMetGluGluGlu 220
QY 601 TTKGARKMTWMTACATGTTGGTGAGAAAGAGACCGGTGGATGGCTACATGAATCAA 660
Db 221 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 240
QY 661 GATGACCTGCCAGAACCGTCGCTCCAGATCATCCGTGACCCCTCCGCAATATAATTCG 720
Db 241 AspAspLeuProArgSerArgSerArgSerValThrLeuProHisIleArg 260
```

```
QY 721 CCAGTGGAGAAATTACAGCAGGAGGTTGGAGAACGTTCTGAGCAATTTCTGAGAGAAG 780
Db 261 ProValGluGluLeuThrGluGluLeuGluValCysSerAsnSerArgGluLys 280
QY 781 ATATATAACCGTTCACCTGGCTCTACTTCTCATCAATGCGTCGAGAAAGCTATTGATACC 840
Db 281 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 300
QY 841 AAAAACHAATCGAGAACCCAGACTCTGGGGGTTTCGAGGCCAGTCTGTGGCCCTGC 900
Db 301 LysThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCys 320
QY 901 CTTTGAACCGTTATCGTGAAGAGTCAGGATGCTCTGTGATCCGAACTGGCATTGC 960
Db 321 LeuArgAsnArgTyrGlyGluGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 340
QY 961 CGCCTTTGTCGAGGAATTCGCAACTGCACTGCTTCTGCCGCGCAGCAGATGCGTGTGCG 1020
Db 341 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 360
QY 1021 ACTGGGCTCTGTGTATTAGCCAAATATCATGCTTGGGAATGTGCATGCTACTTG 1080
Db 361 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 380
QY 1081 AAAAGCCTGAAACAGCAATTTGAAATGCAAGCA 1113
Db 381 LysSerLeuLysGlnGluPheGluMetGlnAla 391

RESULT 9
US-09-925-301-1165
; Sequence 1165, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1165
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1165

Alignment Scores:
Pred. No.: 7,13e-177 Length: 407
Score: 1929.00 Matches: 364
Percent Similarity: 98.11% Conservative: 0
Best Local Similarity: 98.11% Mismatches: 7
Query Match: 98.32% Indels: 0
DB: 9 Gaps: 0

US-10-046-935-2234 (1-1116) x US-09-925-301-1165 (1-407)
QY 1 ATGACGCTCGCGCGTCCGCGAGAAAGATCTCAGAGTAAGAAAGAACTTAAAGAAATTC 60
Db 37 MetAspAlaArgArgValProGlnLysAspLeuArgValLysAsnLeuLysLysPhe 56
QY 61 AGATATGTGAAGTTGATTTCAATGGAACCTCGTCATCCCTGTGATGACAGTTGTGACGC 120
Db 57 ArgTyrValLysLeuIleSerMetGluThrSerSerSerSerSerSerSerSerSerSer 76
QY 121 TTTTCTCTGTATAATTTTGAACACAGCAGGCTGCGAGTTCGGGAGGCTGTAGGACC 180
Db 77 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 96
```

```
QY 181 CCGACGCTGCGAGCACTCTCGACCTCTCAGGGTGGCGATGATGAGTTTCCAGCGCGAGT 240
Db 97 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 116
QY 241 ACCAGGGAGACCAACAAACAAAGCAGAGTCCCGCCAGCCCTCAGAGAAATTTCTGTGACT 300
Db 117 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 136
QY 301 GATTCCAATCCGATTTCAGAGATGAAAGTGAATGAATTTTTTGGAGAAAAGGGCTTTA 360
Db 137 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 156
QY 361 ATATATAACCAACAAACCAATGTTGCAAACTCATGCTCAATAGAAAGCTTCCCT 420
Db 157 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 176
QY 421 GGTCTGTCCTGGGAAGACATCCCTCCAGGCTCCGACTCACAATCAAGAGACACCGCA 480
Db 177 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 196
QY 481 AGCGCTACATTTCCCGGTGTGCTTCCAGGAGAAACCTGAAACGAGAGAGCTGCTCTT 540
Db 197 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgAlaArgProLeu 216
QY 541 ACCAGTCAAGTCCCGGATCCTCGGTCCCTGACGCTCTACCCATGGRGAGAGCCAGT 600
Db 217 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGlu 236
QY 601 TTKGARKMTWMGTACATGTTGGTGAAGAGAGAACCCCTGGATGGCTACATGAATGAA 660
Db 237 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 256
QY 661 GATGACCTGCCAAGAGCGTCTCCAGATCATCCGTGACCCCTTCCGATATAATTCGC 720
Db 257 AspAspLeuProArgSerArgArgSerSerValThrLeuProHisIleLeuArg 276
QY 721 CCAGTGAAGAAATTAACAGAGGAGGTTGGAGAACTCTGACGAAATTTCTCGAGAGAAG 780
Db 277 ProValGluGluIleThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLys 296
QY 781 ATATATAACCGTTCACCTGGGCTCTACTTGTATCAATGCCGTGAGAGAACTATTGATACC 840
Db 297 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 316
QY 841 AAAAACHAATCGCAGAAACCCAGACTCTGGGGCGTTCGAGCCAGTCTGTGGCCCTGC 900
Db 317 LysThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCys 336
QY 901 CTTTGAACCGTTATCGTGAAGAGTCCAGGATGCTCTGTGATCCGAACTGGCATTGC 960
Db 337 LeuArgAsnArgTyrGlyGluGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 356
QY 961 CCGCTTGTGAGGAATCTCAACTGCACTGCTGCGCGCAGCAGAGATGACCGTGTGCG 1020
Db 357 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 376
QY 1021 ACTGGGCTCTGTGTATTAGCCAAATATCATGCTTGGGAATGTGCATGCTACTTG 1080
Db 377 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 396
QY 1081 AAAAGCCTGAAACAGCAATTTGAAATGCAAGCA 1113
Db 397 LysSerLeuLysGlnGluPheGluMetGlnAla 407
```

RESULT 10

US-10-106-698-5654

; Sequence 5654, Application US/10106698

; Publication No. US20030109690A1

; GENERAL INFORMATION:

; APPLICANT: Ruben et al.

; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides

; FILE REFERENCE: PA005P1

; CURRENT APPLICATION NUMBER: US/10/106,698

```
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 5654
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-5654

Alignment Scores:
Pred. No.: 7,13e-177 Length: 407
Score: 1929.00 Matches: 364
Percent Similarity: 98.11% Conservative: 0
Best Local Similarity: 98.11% Mismatches: 7
Query Match: 98.32% Indels: 0
DB: 14 Gaps: 0

US-10-046-935-2234 (1-1116) x US-10-106-698-5654 (1-407)

QY 1 ATGACGCTCCCGCGTCCCGCAGAGATCTCAGAGTAAAGAGAACTTAAGAAATTC 60
Db MetAspAlaArgValProGlnLysAspLeuArgValLysLysAsnLeuLysPhe 56
QY 61 AGATATCTGAGTTGATTCCATCGAAACCTCGTCATCTCTGATGACAGTTGTGACAGC 120
Db 57 ArgTyrValLysLeuIleSerMetGluThrSerSerSerSerSerSerSerSerSerSer 76
QY 121 TTTCTCTCTGATAATTTTGCACACAGAGCTGCAGTCAGTTCCGGAGGCTGTAGACC 180
Db 77 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 96
QY 181 CGCAGCCAGTGCAGGCACTCTGACCTCTCAGGGTGGCGATGAAGTTTCCAGCGCGAGT 240
Db 97 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 116
QY 241 ACCAGGGAGACCAACCAAAAGAGAGAGTCCCGCCAGCCCTCAGAGAAATTCGTGACT 300
Db 117 ThrArgGlyAlaThrAsnLysLysAlaGlnSerArgGlnProSerGluAsnSerValThr 136
QY 301 GATTCCAACCTCCGATTCAGAGATGAAGTGAATGAAATTTTGGAGAAAGGCTTTA 360
Db 137 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 156
QY 361 AATATAAGCAAAACAAAGCAATGCTTGCAGAACTCATGTCTGAATTAGAAAGCTTCCCT 420
Db 157 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 176
QY 421 GCGTCGTTCGTGGAACATCCCTCCAGGCTCCGACTCACAATCAAGAGACCGCGA 480
Db 177 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgProArg 196
QY 481 AGCGGTACATTCCTCGGGTGTGTCTTCAGAGAGAAACCTCAACGGAGAGCTCGCTCTT 540
Db 197 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgArgAlaArgProLeu 216
QY 541 ACCAGGTCAAGTCCCGGATCTCCGGTCCCTTTCAGGCTCTACCCATGGRGAGAGCCAGT 600
Db 217 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGlu 236
QY 601 TTGARKMTWMTGACATGTTGGTGAAGAGAGAGACCGGTGATGCTACATGAATGA 660
Db 237 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 256
QY 661 GATGACCTGCCAGAGCCGCTCCGCTCAGATCATCGTGACCCCTCCGATATATTCGC 720
Db 257 AspAspLeuProArgSerArgArgSerArgSerValThrLeuProHisIleIleArg 276
```


Db 1 MetAspAlaArgArgValProGlnLysAspLeuArgValLysAsnLeuLysPhe 20
QY 61 AGATATGCAAGTTGATTTCATGAACCTCGTCTCATCTCGATGACAGTTGTGACAGC 120
Db 21 ArgTyrValLysLeuLysMetGluThrSerSerSerSerSerSerSerSerSerSer 40
QY 121 TTGTCTTCGATAATTTTCAACACG----- 147
Db 41 PheAlaSerAspAsnPheAlaAsnThrLysProLysPheArgSerAspLysSerGluGlu 60
QY 147 ----- 147
Db 61 LeuAlaSerValPheTyrGluAspSerSerAsnGluSerPheCysGlyPheSerGluSer 80
QY 147 ----- 147
Db 81 GluValGlnAspValLeuAspHisCysGlyPheLeuGlnLysProArgProAspValThr 100
QY 147 ----- 147
Db 101 AsnGluLeuAlaGlyPheHisAlaAspSerAspAspGluSerPheCysGlyPheSer 120
QY 148 -----AGCTGCACTCAGTTCCGGGAAGCTGTAGACCCGC 193
Db 121 GluSerGluIleGlnAspGlyMetArgLeuGlnSerValArgGluGlyCysArgThrArg 140
QY 184 AGCAGTGCAGGCACCTCTGCACCTCTCAGGTCGCGATGAAGTTTCCAGCGCGAGTACC 243
Db 141 SerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSerThr 160
QY 244 AGGGAGCAACCAACAAAAAGCAGAGTCCCGCAGCCCTCAGAGAATTCGTGACTGAT 303
Db 161 ArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThrAsp 180
QY 304 TCCAATCCGATTCCAGAGATGAAGTGAATGAATTTTTCAGAGAAAGGGCTTTAAAT 363
Db 181 SerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeuAsn 200
QY 364 ATAAGCAAAACAAAGCAATGCTTGCAAAACCTCATGTCTGAATAGAAAAGTCCCTGGC 423
Db 201 IleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPheProGly 220
QY 424 TCGTTCGTCGGAAGACATCCCTCCAGGCTCCGACTCACAAATCAAGGACCGCGAAGG 483
Db 221 SerPheArgGlyArgHisProLeuProLeuProGlySerAspSerGlnSerArgArgProArgArg 240
QY 484 CGTACATCCCGGTGTGTGCTTCAGAGAAACCTGACGAGACCTCGTCTCTTACC 543
Db 241 ArgThrPheProGlyValAlaSerArgArgAsnProGluArgArgAlaArgProLeuThr 260
QY 544 AGGTCAAGGTCCCGATCCCTCGGTCTTGCAGCTCTACCCATGCGRAGAGCCAGTTTK 603
Db 261 ArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGluGlu 280
QY 604 GARMTWMTACATGTTGGTGAGAAAGAGGAAGACCGTGGATGGCTCATGAATGAAGAT 663
Db 281 GluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGluAsp 300
QY 664 GACCTGCCAGAACCGTCTCAGATCATCCGTGACCCCTCCGATATATATTCGCCCA 723
Db 301 AspLeuProArgSerArgArgSerSerSerValThrLeuProHisIleArgPro 320
QY 724 GTGGAAGAAATTACAGAGGAGGAGTTGGAGAACGTCTGCACAAATTTCCGAGAGAAGATA 783
Db 321 ValGluGluIleThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLysIle 340
QY 784 TATAACCGTTCACTGGGCTTACTTGTTCATCAATGCCGTCCAGAGACTATTGATACCAAA 843
Db 341 TyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThrLys 360
QY 844 ACAATCTCAGAAACCCAGACTGCTGGCGCTTCGAGCCAGTCTGTGGCCCTTGCCTT 903
Db 361 ThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCysLeu 380

QY 904 CGAAACCGTTATGGTGAAGAGGTGAGGATGCTCTGCTGGATCCGAACCTGGATTCCCGG 963
Db 381 ArgAsnArgTyrGlyGluValArgAspAlaLeuLeuAspProAsnTrpHisCysPro 400
QY 964 CTTGTTCAGGAATCTGCAACTGCACTGCTCTCCGCGCAGCAGATGCGGTGTGCGACT 1023
Db 401 ProCysArgGlyLeCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAlaThr 420
QY 1024 GGGGCTCTGTGTTATTAGCCAAATATCATGGCTTTGGGAATGTGCATCCCTACTGAAA 1083
Db 421 GlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeuLys 440
QY 1084 AGCCTGAAACAGCAATTTGAAATGCAAGCA 1113
Db 441 SerLeuLysGlnGluPheGluMetGlnAla 450

RESULT 12

US-09-851-588-4
; Sequence 4, Application US/09851588
; Patent No. US20020042067A1
; GENERAL INFORMATION:

; APPLICANT: Mack, David
; APPLICANT: Gish, Kurt C.
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER, COMPOSITIONS, AND
; TITLE OF INVENTION: OF SCREENING FOR COLORECTAL CANCER MODULATORS
; FILE REFERENCE: A-68829-1/DJB/JJD/AMS
; CURRENT APPLICATION NUMBER: US/09/851,588
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/642,252
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: US 09/656,002
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-851-588-4

Alignment Scores:
Pred. No.: 1.48e-150 Length: 347
Score: 1656.00 Matches: 338
Percent Similarity: 97.41% Conservative: 0
Best Local Similarity: 97.41% Mismatches: 9
Query Match: 84.40% Indels: 2
DB: Gaps: 0

US-10-046-935-2234 (1-1116) x US-09-851-588-4 (1-347)

QY 1 ATGACGCTCGCGCTGCGCGAGAAAGATCTCAGAGTAAAGAAAGAACTTAAAGAAATTC 60
Db 1 MetAspAlaArgArgValProGlnLysAspLeuArgValLysAsnLeuLysPhe 20
QY 61 AGATATGCAAGTTGATTTCATGGAACCTCGTCTCATCTCTGATGACAGTTGTGACAGC 120
Db 21 ArgTyrValLysLeuLysSerMetGluThrSerSerSerSerSerSerSerSerSer 40
QY 121 TTTGCTTCTGATTAATTTTGCACACGAGGCTGAGTTCGCGGAAGGCTGTAGGACC 180
Db 41 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 60
QY 181 CGCAGCCAGTGCAGGCACCTCTGGACCTCTCAGGGTGGCGATGAAGTTTCCAGCGCGAGT 240
Db 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80
QY 241 ACCAGGGAGCAACCAACAAAAAGCAGAGTCCCGCCAGCCCTCAGAGAAATTCGTGACT 300
Db 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100
QY 301 GATTCACACTCCGATTCCAGAAAGATGAAAGTGAATGAATTTTGGAGAAAAAGGCTTTTA 360


```
Db 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 120
Qy 361 AATATAAGCAAAACAAAGCAATGCTTCCAAACATCATGCTGAATTAAGAGCTCCCT 420
Db 121 AsnLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140
Qy 421 GGCTGTTCCGTGGAGACATCCCTCCAGGCTCCGACTCACAATCAAGGAGACCGGA 480
Db 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160
Qy 481 AGCGGTACATCCCGGTGTTGCTTCCAGAGAAACCTGAAACGAGAGCTCGTCTCTT 540
Db 161 ArgArgThrPheProGlyValAlaSerArgAsnProGluArgAlaArgProLeu 180
Qy 541 ACCAGGTCAAGCTCCGAGTCTCGGGTCCCTTGACGCTTACCCATGGRGAGCCAGT 600
Db 181 ThrArgSerArgSerArgGlyLeuGlySerLeuAspAlaLeuProMetGluGluGlu 200
Qy 601 TTGARKMTWMTACATGTTGGTGAGAAAGAGAACCGTGGATGGCTACATGAATGAA 660
Db 201 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220
Qy 661 GATGACTGCCCAAGCCGCTCGTCAGATCCCGTGCAGTCCGCTTCCGATATATATTCG 720
Db 221 AspAspLeuProArgSerArgArgSerArgSerValThrLeuProHisIleLeuArg 240
Qy 721 CCAGTGAAGAAATACAGA-GGAGGAGTTGGAGAACTGCGAGCAATTCGAG-AGA 778
Db 241 ProValGluGluLeuThrGluGlyValGlyGluArgLeuGlnPheSerLysArg 260
Qy 779 AGATATATAACCGTCACTGGCTTACTTGTCTCAATGCGTCAAGAGACTATTCATA 838
Db 261 ArgTyrIleThrValHisTyrAlaLeuLeuValIleAsnAlaValArgLeuLeuLe 280
Qy 839 CCAAAACAACTGCAGAAACCCAGACTGCTGGGGCGTTGAGGCCAGTTCGTGGCCCT 898
Db 281 ProLysGlnThrAlaGluThrGlnThrAlaGlyAlaPheGluAlaSerSerValAlaPro 300
Qy 899 GCCTTCCAAACCGTATGTTGTAAGAGTCAAGGATGCTCTGCTGGATCCGAACTGGATT 958
Db 301 AlaPheGluThrValMetValLysArgSerGlyMetLeuCysTrpIleArgThrGlyLe 320
Qy 959 GCGCGCTTGTGAGGAATTCGAACCTGCAGTTCCTGCGCGAGCGAGATGCGAGGTGTG 1018
Db 321 AlaArgLeuValGluGluSerAlaThrAlaValSerAlaGlySerGluMetAspGlyVal 340
Qy 1019 CGACTGGGTCTGTGTATT 1039
Db 341 ArgLeuGlySerLeuCysIle 347
```

RESULT 13

```
US-10-295-027-1194
; Sequence 1194, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCES: 018501-01250005
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
```

```
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1194
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1194

Alignment Scores:
Pred. No.: 1,48e-150 Length: 347
Score: 1656.00 Matches: 338
Percent Similarity: 97.41% Conservative: 0
Best Local Similarity: 97.41% Mismatches: 9
Query Match: 84.40% Indels: 2
DB: 14 Gaps: 0

US-10-046-935-2234 (1-1116) x US-10-295-027-1194 (1-347)
Qy 1 ATGACGCTCCGCGCTGCCGCGAAGAGTCTCAGAGTAAGAGAACTTAAGAAATTC 60
Db 1 MetAspAlaArgArgValProGlnLysAspLeuArgValLysLysLysPhe 20
Qy 61 AGATATGTAAGTTGATTTCCATGGAACCTGTCATCCTCTGATGACAGTCTGACAGC 120
Db 21 ArgTyrValLysLeuLeuSerMetGluThrSerSerSerSerSerSerSerSer 40
Qy 121 TTTGCTTCTGATATTTTGCACACACAGAGCTGCAGTCAGTTCGGGAAGGCTGTAGACC 180
Db 41 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 60
Qy 181 CGCAGCCAGTGCAGGCACCTCTGGACCTCTCAGGTCGCATGAGTTTCAGCGCGGAGT 240
Db 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80
Qy 241 ACCAGGGAGCAACCAACAAAAAGCAGAGTCCCGCCAGCCCTCAGAGATTCGTGACT 300
Db 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100
Qy 301 GATTCACATCCGATTTCAGAAGATGAAAGTGGAAATGAAATTTTGGAGAAAAGGGCTTTA 360
Db 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 120
Qy 361 AATATAAGCAAAACAAAGCAATGCTTCCAAACATCATGCTGAATTAAGAGCTTCCT 420
Db 121 AsnLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140
Qy 421 GGCTGTTCCGTGGAGACATCCCTCCAGGCTCCGACTCACAATCAAGGAGACCGGA 480
Db 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160
Qy 481 AGCGGTACATCCCGGTGTTGCTTCCAGAGAAACCTGAAACGAGAGCTCGTCTCTT 540
Db 161 ArgArgThrPheProGlyValAlaSerArgAsnProGluArgAlaArgProLeu 180
Qy 541 ACCAGGTCAAGCTCCGAGTCTCGGGTCCCTTGACGCTTACCCATGGRGAGCCAGT 600
Db 181 ThrArgSerArgSerArgGlyLeuGlySerLeuAspAlaLeuProMetGluGluGlu 200
Qy 601 TTGARKMTWMTACATGTTGGTGAGAAAGAGAACCGTGGATGGCTACATGAATGAA 660
Db 201 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220
Qy 661 GATGACTGCCCAAGCCGCTCGTCAGATCCCGTGCAGTCCGCTTCCGATATATATTCG 720
Db 221 AspAspLeuProArgSerArgArgSerArgSerValThrLeuProHisIleLeuArg 240
Qy 721 CCAGTGAAGAAATACAGA-GGAGGAGTTGGAGAACTGCGAGCAATTCGAG-AGA 778
Db 241 ProValGluGluLeuThrGluGlyValGlyGluArgLeuGlnPheSerLysArg 260
Qy 779 AGATATATAACCGTCACTGGCTTACTTGTCTCAATGCGTCAAGAGACTATTCATA 838
Db 261 ArgTyrIleThrValHisTyrAlaLeuLeuValIleAsnAlaValArgLeuLeuLe 280
Qy 839 CCAAAACAACTGCAGAAACCCAGACTGCTGGGGCGTTGAGGCCAGTTCGTGGCCCT 898
Db 281 ProLysGlnThrAlaGluThrGlnThrAlaGlyAlaPheGluAlaSerSerValAlaPro 300
Qy 899 GCCTTCCAAACCGTATGTTGTAAGAGTCAAGGATGCTCTGCTGGATCCGAACTGGATT 958
Db 301 AlaPheGluThrValMetValLysArgSerGlyMetLeuCysTrpIleArgThrGlyLe 320
Qy 959 GCGCGCTTGTGAGGAATTCGAACCTGCAGTTCCTGCGCGAGCGAGATGCGAGGTGTG 1018
Db 321 AlaArgLeuValGluGluSerAlaThrAlaValSerAlaGlySerGluMetAspGlyVal 340
Qy 1019 CGACTGGGTCTGTGTATT 1039
Db 341 ArgLeuGlySerLeuCysIle 347
```


82	ATGGAACCTCGTCACTCCTGATGACAGTTGTGACAGCTTTGCTCTGATAATTTTGC	141	QY
1	MetGluThrSerSerSerAspAspSerCysAspSerPheAlaSerAspAsnPheAla	20	Db
142	AACAGAGGCTGCAGTCAGTTCGGGAAGGCTGTAGACCCGCGAGCCAGTCAGGCATCTCT	201	QY
21	AsnThrArgLeuGlnSerValArgGluGlyCysArgThrArgSerGlnCysArgHisSer	40	Db
202	GGACTCTTCAGGGTGGCGATCAAGATTTCCAGCGCGGAGTACCAGGGGAGCAACCAACA	261	QY
41	GlyProLeuArgValAlaMetLysPheProAlaArgSerThrArgGlyAlaThrAsnLys	60	Db
262	AAAGCAGAGTCCCGCCAGCCCTCAGAGAATCTGTGACTGATTCMACTCCGATTCAGAA	321	QY
61	LysAlaGluSerArgGlnProSerGluAsnSerValThrAspSerAsnSerAspSerGlu	80	Db
322	GATCAAAAGTGGAAATGAAATTTTTGGAGAAAAGGCTTTAAATAAAGCAACAAACGA	381	QY
81	AspGluSerGlyMetAsnPheLeuGluLysArgAlaLeuAsnIleLysGlnAsnLysAla	100	Db
382	ATGCTTCGAAACCTCATGTCTGAATTAGAAAGCTTCCTCGGCTCGTTCGCTGGAGACAT	441	QY
101	MetLeuAlaLysLeuMetSerGluLeuGluSerPheProGlySerPheArgGlyArgHis	120	Db
442	CCCCCTCCAGGCTCCGACTCACAAATCAAGAGACCCGAGGCGGTACATCCCGGGTGTT	501	QY
121	ProLeuProGlySerAspSerGlnSerArgArgProArgArgThrPheProGlyVal	140	Db
502	GCTTCCAGGAGAAACCTCGAACGAGAGCTCGTCCTCTTACACAGGTCGAGGTCGCCGATC	561	QY
141	AlaSerArgArgAsnProGluArgArgAlaArgProLeuThrArgSerArgSerArgIle	160	Db
562	CTCGGGTCCCTTCACGCTCTACCCATGGRGAGAGCCAGTTTAKARNTWWTACATGTTG	621	QY

Blank

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 12, 2004, 15:07:46 ; Search time 23 Seconds
(without alignments)
6435.729 Million cell updates/sec

Title: US-10-046-935-2234
Perfect score: 1962
Sequence: 1 atggacgctgcgcgctgcc.....aattgaaatgcaagcataa 1116

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 956278

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2_1/USPRO.spool_P/US10046935/runat_12102004_155226_15267/app_query.fasta_1.1287
-DB=Issued_Patents_AA -QPMI=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10046935 @CIGN 1.1 37 @runat 12102004_155226_15267 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pcp:*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pcp:*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pcp:*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pcp:*
5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pcp:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1656	84.4	347	4	US-09-851-588-4
2	1520.5	77.5	320	4	US-09-851-588-2
3	125.5	6.4	433	4	US-09-252-991A-20376
4	125	6.4	754	4	US-09-976-594-375
5	116.5	5.9	904	4	US-09-976-594-615
6	110	5.6	536	3	US-09-188-930-185
7	110	5.6	536	4	US-09-312-283C-185
8	110	5.6	742	3	US-08-791-115B-5
9	107	5.5	380	4	US-09-270-767-43626
10	106	5.3	582	4	US-09-252-991A-21510
11	106	5.4	820	4	US-09-248-796A-17231
12	105	5.4	308	4	US-09-252-991A-31991

13	105	5.4	679	4	US-09-252-991A-27111	Sequence 27111, A
14	103.5	5.3	947	4	US-09-418-780A-1	Sequence 1, Appli
15	103.5	5.3	947	4	US-09-392-714-23	Sequence 23, Appli
16	102.5	5.2	499	4	US-09-252-991A-30259	Sequence 30259, A
17	102	5.2	266	4	US-09-252-991A-32512	Sequence 32512, A
18	102	5.2	301	4	US-09-248-796A-18853	Sequence 18853, A
19	102	5.2	310	4	US-09-270-767-42373	Sequence 42373, A
20	102	5.2	395	4	US-09-252-991A-30497	Sequence 30497, A
21	101.5	5.2	944	3	US-09-449-285A-2	Sequence 2, Appli
22	101.5	5.2	1850	4	US-09-620-093A-5	Sequence 5, Appli
23	100.5	5.1	3075	2	US-08-460-309-5	Sequence 5, Appli
24	100.5	5.1	3075	2	US-08-125-077-5	Sequence 5, Appli
25	100	5.1	559	1	US-08-320-559-31	Sequence 31, Appli
26	100	5.1	559	3	US-08-545-860D-31	Sequence 31, Appli
27	100	5.1	559	3	US-09-538-092-1261	Sequence 1261, Ap
28	100	5.1	559	5	PCT-US94-04496-31	Sequence 31, Appli
29	99	5.0	667	4	US-09-248-796A-19663	Sequence 19663, A
30	99	5.0	2200	4	US-09-796-575-2	Sequence 2, Appli
31	98.5	5.0	295	3	US-09-199-637A-341	Sequence 341, App
32	97.5	5.0	357	4	US-09-252-991A-20112	Sequence 20112, A
33	97.5	5.0	454	2	US-09-159-385-1	Sequence 1, Appli
34	97.5	5.0	454	3	US-09-186-277-1	Sequence 1, Appli
35	97.5	5.0	590	4	US-09-312-283C-409	Sequence 409, App
36	97.5	5.0	786	4	US-09-509-802-2	Sequence 2, Appli
37	97.5	5.0	787	3	US-09-188-930-334	Sequence 334, App
38	97.5	5.0	787	4	US-09-312-283C-334	Sequence 334, App
39	97.5	5.0	1342	2	US-08-484-438-9	Sequence 9, Appli
40	97.5	5.0	1523	4	US-09-538-092-955	Sequence 955, App
41	97	4.9	2802	3	US-09-542-331-1	Sequence 1, Appli
42	97	4.9	2802	3	US-09-510-791-1	Sequence 1, Appli
43	96.5	4.9	439	4	US-09-252-991A-21361	Sequence 21361, A
44	96.5	4.9	1953	4	US-09-917-254-92	Sequence 92, Appl
45	96	4.9	494	4	US-09-538-092-1282	Sequence 1282, Ap

ALIGNMENTS

RESULT 1
; Sequence 4, Application US/09851588
; Patent No. 6682890
; GENERAL INFORMATION:
; APPLICANT: Mack, David
; APPLICANT: Gish, Kurt C.
; APPLICANT: Wilson, Keith E.
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER, COMPOSITIONS, AND N
; FILE REFERENCE: A-68829-1/DJB/JJD/AMS
; CURRENT APPLICATION NUMBER: US/09/851,588
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/642,252
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: US 09/656,002
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-851-588-4

Alignment Scores:
Pred. No.: 9.2e-177 Length: 347
Score: 1656.00 Matches: 338
Percent Similarity: 97.41% Conservative: 0
Best Local Similarity: 97.41% Mismatches: 9
Query Match: 84.40% Indels: 2
DB: 4 Gaps: 0

US-10-046-935-2234 (1-1116) x US-09-851-588-4 (1-347)

QY 1 ATGGACGCTGCGCGCTGCGCAGAGAAGATCTCAGAGTAAGAAGAACTTAAAGAATTC 60

```
Db 1 MetAspAlaArgArgValProGlnLysAspLeuArgValLysAsnLeuLysLysPhe 20
QY 61 AGATATGTGAAGTTGATTCATCGGAAACCTCGTCATCTCTGATCAGACTTGTGACAGC 120
Db 21 ArgTyrValLysLeuIleSerMetGluThrSerSerSerSerSerSerSerSerSer 40
QY 121 TTTCCTCTGATATTTTGCACACAGAGCTCGAGTCACTCGGGAAGCTCTAGGACC 180
Db 41 PheAlaSerAspAsnPheAlaAsnThrArgLysGlnSerValArgGluGlyCysArgThr 60
QY 181 CGCAGCAGTCAGGCACTCTGACCTCTCAGGTTGCGGATGAAGTTTCAGCGCGAGT 240
Db 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80
QY 241 ACCAGGGGACACCAACAAAGAGAGAGTCCGCCAGCCCTCAGAGAATTTCTGTGACT 300
Db 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100
QY 301 GATTCCAACTCCGATTCCAGAGATGAAGTGGAAATGAATTTTGGAGAAAAGGCTTTA 360
Db 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 120
QY 361 AATATAAGCAAAACAAAGCAATCTCTGCAAACTCATGTCTGAATTAGAAAGCTTCCCT 420
Db 121 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140
QY 421 GGCTCGTTCGTGGAGACATCCCTCCAGGCTCCGACTCACAAATCAAGAGACCCGGA 480
Db 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160
QY 481 AGCGGTACATTCGCGGTGTGTCTCCAGAGAAACCTGAAACGAGAGCTCGTCCCTTT 540
Db 161 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgAlaArgProLeu 180
QY 541 ACCAGGTCAAGGTCCCGATCTCTCGGTCCCTTGACCGCTCTACCCATGCGRAGAGCCAGT 600
Db 181 ThrArgSerArgSerArgLysLeuLysSerLeuAspAlaLeuProMetGluGluGlu 200
QY 601 TTKGARKMTWGTACNTGTTGGTGGAGAAAGGAGACCGTGGATGCTCATCAATGAAGAA 660
Db 201 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220
QY 661 GATGACCTGCCAGAACCGTCGCTCCAGATCATCCGTGACCCCTCCGCTATATATTCGC 720
Db 221 AspAspLeuProArgSerArgSerArgSerValThrLeuProHisIleLeuArg 240
QY 721 CCAGTGAAGAAATTAAGA- GGAGGAGTTGGAGAACGTCCTGCGAGCAATTTCTCGAG-AGA 778
Db 241 ProValGluGluIleThrGluGlyGlyValGlyLysArgLysThrValAspGlyTyrMetAsnGlu 260
QY 779 AGATATATAACCGTTCACCTGGCTCTACTTGTCTCATCAATGCGTCAGAAGACTATTGATA 838
Db 261 ArgTyrIleThrValHisIleProAlaLeuLeuValIleAsnAlaValArgLeuLeuIle 280
QY 839 CCAAAAACAACTGACAGAAACCCAGACTGCTGGGCGTTTCGAGGCCAGTTCGTGGCCCTC 898
Db 281 ProlLysGlnThrAlaGluThrGlnThrAlaGlyAlaPheGluAlaSerSerValAlaPro 300
QY 899 GCCTTCGAACCGTTATGTTGAAGAGTCAGGATGCTCTCTGATCCGAACTGGCATTT 958
Db 301 AlaPheGluThrValMetValLysArgSerGlyMetLeuCysTrpIleArgThrGlyLe 320
QY 959 GCCCGCTTGTGAGGAGATCTGCAACTGCACTGTCAGTTCTGCGCGCAGCAGATGCGGTGTG 1018
Db 321 AlaArgLeuValGluGluSerAlaThrAlaValSerAlaGlySerGluMetAspGlyVal 340
QY 1019 CGACTGGGGTCTCTGTGTATT 1039
Db 341 ArgLeuGlySerLeuCysIle 347
```

RESULT 2

US-09-851-588-2

```
; Sequence 2, Application US/09851588
; Patent No. 6682890
; GENERAL INFORMATION:
; APPLICANT: Mack, David
; APPLICANT: Gish, Kurt C.
; APPLICANT: Wilson, Keith B.
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER, COMPOSITIONS, AND
; TITLE OF INVENTION: OF SCREENING FOR COLORECTAL CANCER MODULATORS
; FILE REFERENCE: A-68829-1/DJB/JJD/AMS
; CURRENT APPLICATION NUMBER: US/09/851,588
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/642,252
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: US 09/656,002
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-851-588-2
```

```
Alignment Scores:
Pred. No.: 1,44e-161 Length: 320
Score: 1520.50 Matches: 306
Percent Similarity: 96.26% Conservative: 3
Best Local Similarity: 95.33% Mismatches: 10
Query Match: 77.50% Indels: 2
DB: 4 Gaps: 1
```

US-10-046-935-2234 (1-1116) x US-09-851-588-2 (1-320)

```
QY 82 ATGMAACCTCGTCATCTCTGATGACAGTTGTGACAGTTGTCTGATAATTTTGA 141
Db 1 MetGluThrSerSerSerSerSerSerSerSerSerSerPheAlaSerAspAsnPheAla 20
QY 142 AACACGAGGTCCAGTCAGTTCGGGAAAGGTGTAGGACCCGCGAGCCAGTCAGGCACCTCT 201
Db 21 AsnThrArgLeuGlnSerValArgGluGlyCysArgThrArgSerGlnCysArgHisSer 40
QY 202 GGACCTCTCAGGTGGCGATGAAATTTCCAGCGGAGTAGTACCAGGGAGCAACCAACAA 261
Db 41 GlyProLeuArgValAlaMetLysPheProAlaArgSerThrArgGlyAlaThrAsnLys 60
QY 262 AAACGAGGTCCCGCCAGCCCTCAGAGAATTTCTGTGACTGATTCCTCAACTCCGATTCAGAA 321
Db 61 LysAlaGluSerArgGlnProSerGluAsnSerValThrAspSerAsnSerAspSerGlu 80
QY 322 GATGAAAGTGAATGAATTTTGGAGAAAGGGCTTTAAATATAAGCAACCAACAAAGCA 381
Db 81 AspGluSerGlyMetAsnPheLeuGluLysArgAlaLeuAsnIleLysGlnAsnLysAla 100
QY 382 ATGCTTCGAAACCTCATGCTCGAATTAGAAGCTTCCCTGCTGCTCCGTGGAAGACAT 441
Db 101 MetLeuAlaLysLeuMetSerGluLeuGluSerPheProGlySerPheArgGlyArgHis 120
QY 442 CCCCTCCAGGCTCCGACTCACATCAAGGAGACCCGCGAGCGGTACATTCCTCCGGTGT 501
Db 121 ProlLeuGlySerAspSerGlnSerArgArgProArgArgThrPheProGlyVal 140
QY 502 GTTTCAGGAGAAACCTCGAACCGGAGAGTCGTCCTCTTACCAGGTCAAGTCCCGGATC 561
Db 141 AlaSerArgArgAsnProGluArgAlaArgProLeuThrArgSerArgSerArgIle 160
QY 562 CTCGGGTCCCTTGACGCTCTACCATGCGRAGAGCCAGTTTGTGARKMTWGTACATGTTG 621
Db 161 LeuGlySerLeuAspAlaLeuProMetGluGluGluGluGluAspLysTyrMetLeu 180
QY 622 GTGAGAAAGAGAGAACCGGTGGATGGCTTACATGAATGAAGATGACCTCCAGCAAGCCGT 681
Db 181 ValArgLysArgLysThrValAspGlyTyrMetAsnGluAspAspLeuProArgThrArg 200
```

```
QY 682 CGCTCCAGATCATCCGTGACCCCTTCCGATATTAATTCGCCAGTGGAGAAATTACAGAG 741
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 201 ArgTyrArgSerValThrLeuProHisIleIleArgProValGluGluIle-GlnLy 220

QY 742 G---AGAGTTGGAGAACGCTGCGCAAAATTCGAGAGAAGATATATAACCGTTCACTG 798
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 220 sGluArgSerTrpArgThrSerAlaAlaIleLeuGluGluLysIleIleThrValHisTr 240

QY 799 GGCTCTACTTGTCTATCATCCGCTCAGAGACTATTGATACCAAAACAAATCGCAGAAC 858
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 240 pAlaLeuLeuValIleAsnAlaValArgLeuLeuIleProLysGlnThrAlaGluTh 260

QY 859 CCAGACTGCTGGGGCGTTCCAGGCGCAGTCTGTGGCCCTCGCTTCGAAACCGTTATGGT 918
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 260 rGlnThrAlaGlyAlaPheGluAlaSerSerValAlaProAlaPheGluThrValMetVa 280

QY 919 GAAGAGTTCAGGAGTGTCTGTGCGATCCGAACCTGCGCATTCGCCCTTTCGAGGAATC 978
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 280 lLysArgSerGlyMetLeuCysTrpIleArgThrGlyLeAlaArgLeuValGluGluSe 300

QY 979 TGCRACTGCAGTTTCTCCGCGCAGGAGATGCGAGGTGTCGACTGGGTCTGTGTAT 1038
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 300 rAlaThrAlaValSerAlaGlySerGluMetAspGlyValArgLeuGlySerLeuCysIl 320

QY 1039 T 1039
|
Db 320 e 320
```

RESULT 3

```
US-09-252-991A-20376
; Sequence 20376, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20376
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20376
```

```
Alignment Scores:
Pred. No.: 7.52e-05 Length: 433
Score: 125.50 Matches: 68
Percent Similarity: 31.34% Conservative: 16
Best Local Similarity: 25.37% Mismatches: 76
Query Match: 6.40% Indels: 109
DB: 4 Gaps: 14
```

```
US-10-046-935-2234 (1-1116) x US-09-252-991A-20376 (1-433)
```

```
QY 343 TTGGAGAAAAGGGCTTTAAATATATAAGCAAAACAAAGCAATGCTTGCAAAACTCATGTCT 402
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2 LeuLysAlaLeuAlaSerValMetArgAsnThrProGlySerLeuAlaGlyThrPheSer 21

QY 403 GAATTA-----GAAAGCTTCCTGGCTCGTTCGTTGGAAGACATCCCTCCCA 450
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 22 SerLeuProValProSerArgGlyPheProGlySerArgAlaAlaArg---ProPhePro 40

QY 451 GGCTCCGACTCACAA-----TCAGG 471
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 41 AlaSerAspSerArgMetProSerGlyTrpProValProCysAlaGlyThrProArgArg 60

QY 472 AGACCGCGAAGCGGTACATTCGCCGGGTGTTGCTTCCAGGAGAAACCTCGAAGCGGAGACT 531
```

```
Db 61 SerThrArgArgSerAlaAlaGlySerAlaIleArgArgThrProAlaGlySerVal 80
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 532 GTCCTCTTT---ACCAGGTCAAGTCCCGATCCTCGGTCCCTTGGAGCTCTACCCATG 588
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 81 GlyProCysArgCysGluProArg-ProSerAlaAlaGlyProThrSer-----ProTr 98
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 589 GRGAGAGCCAGTTTGARKMTMGTACATCTGTGTGAGAAAGAGAACCGTGGTGGC 648
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 98 pProAlaPro-----ArgAlaIleProAlaArgAlaAla 109
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 649 TACATGAATGAAGATGACCTGCCAGAAAGCGTCCGTCCAGATCATCCGTGACCCCTCCG 708
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 109 ----- 109

QY 709 CATATAATTCGCCAGTGGAGAAATACAGAGGAGAGTGGAGAACGTC----- 759
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 110 -----GlyArgThrGlyArgSerTrpProSerProAlaGlu 123
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 760 -TGCAGCAATTCGAGAGAAGATATATAACCGTTCACTGGGCTCTACTTGTTCATCAATG 818
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 123 yCysSerAlaAlaArg-----SerArgProThrArgCy 135
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 819 CCGTGAGAGACTATTGATACCAAAACAAACTGCAGAAACCCAGACTGCTGGGGCGTTCG 878
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 135 sArgSerArgThrIleArgGlyGluThrGlyAsnTrpSerProThrAlaGlySerAlaAr 155
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 879 A-----GGCAGTTCTGTGGCCCTCGCTTCGAAACCGTTATGTTGGAAGAGTTCAGGA 932
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 155 gThrGlyProThrCysCysGlyProCys-----ArgTyr 166
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 933 TGCTCTGCTGATCCGAACCTGGCATTCGCCCG----- 963
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 167 -----TrpArgTrpProPheArgProGlyThrAlaGlyThrProAr 180
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 964 ----CCTTGTGAGGAATCTGCAACTGC----- 987
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 180 gSerProCysArgGlyProGlySerCysCysAsnArgAlaArgLeuProGlySerGlyAr 200
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 988 ----AGTTTCTGCGCGCAGCGA 1005
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 200 gGlySerAlaCysArgSerArg 207
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4
US-09-976-594-375
; Sequence 375, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 375
; LENGTH: 754
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 3688791CD1
US-09-976-594-375

Alignment Scores:
Pred. No.: 0.00012 Length: 754
Score: 125.00 Matches: 75
Percent Similarity: 38.48% Conservative: 52
Best Local Similarity: 22.73% Mismatches: 145
Query Match: 6.37% Indels: 58
```

```
DB: 4 Gaps: 12
US-10-046-935-2234 (1-1116) x US-09-976-594-375 (1-754)
QY 4 GAGCTCGCGCGTGGCGGAGAAAGATCTCAGAGTAAAGAACTTAAAGAAATTCAGA 63
Db 437 AspIleArgAsnSerGluLysGluAspLysTyrLysAsnLysValLysLysArgAla 456
QY 64 TATGTGAAGTTGATTTTCATGGAACCTCGCTCATCTCTCATGCACAGTTGTGACAGCTTT 123
Db 457 LysSerLysSerArgSerLysLysGluLysSerLysSerLysGluArgAspSerLys 476
QY 124 GCTTCTGATTAATTTGCAACACAGAGCTGCAGTTCGGGAAGCTGTAGACCCGC 183
Db 477 HisAsnArgAsn-----GluLysLysArgMetArg 486
QY 184 AGCAGTGCAGGCACTCTGACCTCTCAGGTTGGCGATGAAGTTTCAGCGCGGAGTACC 243
Db 487 SerArgSerLysGlyArgAspHisGluAsnValLysGluLysGluLysGlnSerAspSer 506
QY 244 AGGGGAGCAACCAACAAAGACAGAGTCCCGCCAGCCCTCAGAGAAATTCGTGACTGAT 303
Db 507 LysGlyLysAspGlnGluArgSerArgSerLysGluLysSerLysGln-----LeuGlu 524
QY 304 TCCAACTCCGATTCAGAAGATGAAGTGAATGAATTTTGGAGAAAGCGCTTTAAAT 363
Db 525 SerLysSerAsnGluHisAspHisSerLysSerLysGluLysAspArgAla----- 542
QY 364 ATAAACCAACAAACCAATGCTTGCNAACACTCATGTCGAATTAGAAAGCTTCCCTGGC 423
Db 543 -----GlnSerArgSerArgGluCysAspIleThrLysGlyLysHisSerTyrAsnSer 560
QY 424 TCGTTCCTCGTGGAAAGACATCCCTCCAGGCTCCGACTCACAATCAAGAGACCGCGAAGG 483
Db 561 ArgThrArgGluArg-----SerArgSerArgAspArgSerArgValArgSer 577
QY 484 CGTACATTCGCGGTGTGCTCCAGG-----AGAAACCTCGAA 522
Db 578 ArgThrHisAspArgAspArgSerArgSerLysGluLysArgTyrArgGluGlnGlu 597
QY 523 CGGAGAGCTCGTCTCTTACAGGTCAAGTCCGAGTCCGAGTCTCGGGTCCCTTGACGCTCA 582
Db 598 TyrArgArgArgGlyArgSerArgSerArgSerArgGluArg-----ArgThr 611
QY 583 CCCATGGGAGACCGAGTATTTKGARKMTWGTACATGTTGTGTGAGAAAGAGAACCGGTG 642
Db 612 ProGlyArgSerArgSerLysAsp-----ArgArgArgArgArgArg 626
QY 643 GATGGCTACATGAATGAGATGACCTGCCAGAGACCGTCTGCCATCATCGTGACC 702
Db 627 AspSerArgSerGluArgGluGluSerGlnSerArgAsnLysAspLys----- 643
QY 703 CTTCCGCATATAATTCGCCAGTGGAGAAATACAGAGGAGGAGTGGAGAGCTGTCG 762
Db 644 -----TyrArgAsnGlnGluSerLysSerHisArgLysGluAsn----- 657
QY 763 AGCAATTCGAGAGATATATAACCGTTCATCTGGGC---TCTACTTGTCTCAATGTC 819
Db 658 SerGluSerGluLysArgMetTyrSerLysSerArgAspHisAsnSerAsnSer 677
QY 820 CGTCAGAGACTATTGAT-----ACCAAAACAACACTCCAGAACCCA 861
Db 678 ArgGluLysLysAlaAspArgAspGlnSerProPheSerLysLysLysGlnSerGln 697
QY 862 GACTGCTGGGCGGTTCGAGGCGAGTCTGTGGCCCTCGCTCGAAACCGTTATGGTGA 921
Db 698 Asp-----AspGluLysSerSerMetLeuLysAsnLysGluAspGlu 712
QY 922 GAGGTACAGGATGCTCTGCTGGATCCGAAC 951
Db 713 LysIleArgSerValGluLysGluAsn 722
```

RESULT 5

```
US-09-976-594-615
; Sequence 615, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 615
; LENGTH: 904
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; FEATURE:
; OTHER INFORMATION: Incyte ID No. 6673549 2789525CD1
US-09-976-594-615
```

```
Alignment Scores:
Pred. No.: 0.00121 Length: 904
Score: 116.50 Matches: 65
Percent Similarity: 37.60% Conservative: 29
Best Local Similarity: 26.00% Mismatches: 111
Query Match: 5.94% Indels: 45
DB: 4 Gaps: 11
```

```
US-10-046-935-2234 (1-1116) x US-09-976-594-615 (1-904)
QY 16 GTCCGCGAGAAAGATCTCAGAGTAAAGAAAGAACTTAAAGAAATTCAGATATGTG----- 69
Db 119 IleProSerAlaPheLeuGluLeuLysGluGluLysGlnArgGlnIleGluGln 138
QY 70 ---AAGTTGATTTCCATGGAAACCTCGTCACTCTCATCTCATGACAGTGTGTGACAGCTTTGCT 126
Db 139 GluLysLeuAlaSerMetLysLysGlnAspGluAspLysAspLysArgAspLysGluGlu 158
QY 127 TCTGATAATTTCCAAACACGAGCTGCAGTCA-----GTTCCGGAAGCTGTAGG 177
Db 159 LysGluSerArgGluLysArgGluArgSerArgSerProArgArgLysSerArg 178
QY 178 ACCCGCAGCCAGTGCAGGCACTCTCGACCTCTCAGGTCGCGATGAAG---TTCCACAGC 234
Db 179 SerProSerProArgArgArgSerSerProValArgArgGluArgLysArgSerHisSer 198
QY 235 CGAGTACCGAGGAGCAACCAACAAAGACAGAGTCCCGCAGCCCTCAGAGAAATTC 294
Db 199 ArgSerProArgHisArgThrLysSerArgSerProSerProAlaProGluLysLysGlu 218
QY 295 GTCACTGATTCACCACTCCGATTCCAGAGATGAAGTGAATGATTTTTCGAGAAAGG 354
Db 219 LysThrProGlu-----LeuProGluPro 226
QY 355 GCTTTAAATATAAGCAAAACAAAGCAATGCTTGCAAAACATCATGTCTGAATTA----- 408
Db 227 SerValLysValLysGluProSerValGlnGluAlaThrSerThrSerAspIleLeuLys 246
QY 409 -----GAAAGCTTCCTCGCTCGTCCGTGGAAGACATCCCTCCAGGCTCC 456
Db 247 ValProLysProGluProLysPro-----GluProLysGluProSerProGluLys 263
QY 457 GACTCTCAATCAAGGAGACCGCGAAGCGGTACATTCCCGGTGTGCTTCCAG----- 510
Db 264 AsnSerLysLysGluLysGluLysArgProArgSerArgSerArgSerArgSerLys 283
QY 511 -----AGAAACCTTGAACCGGAGAGCTCGTCTCTTACCAGGTCA 549
Db 284 SerArgSerArgThrArgSerArgSerProSer---HisThrArgProArgArgHis 302
```



```

QY 550 AGGTCCTCGGATCTCTGGGTCCTTACGCTCTACCCATGGRGAGACGAGTTCGARKMT 609
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
303 ArgSerArg-----SerArgSerTyrSerProArgArgProSerProArgArg 319
QY 610 MWGTACATGTTGGTCAGAAAGAGGAGACCGTGGATGGCTACATGAATGAAGATGACCTG 669
Db ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
320 ArgProSer-----ProArgArgThrProProArgArgMetProProPro----- 335
QY 670 CCCAGAACCGTCGCTCCAGATCATCCGTG 699
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
336 ProArgHisArgArgSerArgSerProVal 345

RESULT 6
US-09-188-930-185
; Sequence 185, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/188.930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 185
; LENGTH: 536
; TYPE: PRT
; ORGANISM: mouse
US-09-188-930-185

Alignment Scores:
Pred. No.: 0.00472 Length: 536
Score: 110.00 Matches: 61
Percent Similarity: 37.28% Conservative: 43
Best Local Similarity: 21.86% Mismatches: 111
Query Match: 5.61% Indels: 64
DB: 3 Gaps: 12

US-10-046-935-2234 (1-1116) x US-09-188-930-185 (1-536)
QY 172 TGTAGACCCGC-----ACCCAGTCGAGCAC 198
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
250 CysArgProArgProArgAlaCysAlaSerLeuLeuLeuMetGlnArgCysTrpHis 269
QY 199 TCTGGACCT---CTCAGGTGGCGATGAAGTTTCCAGCGCGAGTACCAGGGGAGCAACC 255
Db ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
270 AlaAspProGlnValArgProThrPheGlnGluLeuThrSerGluThrGluAspLeuCys 289
QY 256 AACAAACACACAG-----TCCGCCAGCCCTCAGAGAATCTGTGACT 300
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
290 GluLeuProAspGluGluValLysAspLeuAlaHisGluProGlyGluLysSerSerLeu 309
QY 301 GATTCAACCTCCGATTCAGAGATGAAGTGAATGTAATTTTGGAGAAAGGGCTTTA 360
Db ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
310 GluSerLysSerGluAlaArgProGluSerSerArg-----LeuLysArgAlaSerAla 327
QY 361 AATATAAGCAACAAAGCAATGCTTGCACAACTCATGTCTGAATTAGAAAGC----- 414
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
328 ProProPheAspAsnAspCysSerLeuSerGluLeuLeuSerGlnLeuAspSerGlyLe 347
QY 415 TTCCCTGGCTGCTCCGTGGAAGACATCCCTCCAGGCTCCGACTCACAATCAAGGAGA 474
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
348 PheArgLeuLeuLysGlyProGluGluLeuSerArgSerSerGluCysLysLeu 367
QY 475 CCCGGAAGG-----CGTACATTCGCGGTGTCCTCCAGGAGAAACCTGACGG 525
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
368 ProSerSerSerSerGlyLysArgLeuSerGlyValSerSerValAspSerAlaPheSer 387

```

```

QY 526 AGAGTCGTCTCTTACAGGTCCAGGTCCTCGGATCTCTCGGTCCTTACGCTCTACCC 585
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
388 SerArgGlySerLeuSerLeuSerPheGluArgGluAlaSerThrGlyAspLeuGlyPro 407
QY 586 ATGGRGAGAGCCAGTTTGGARKMTMWGTACATGTTGGTGAGAAAGAGGAGACCGCTGAT 645
Db ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
408 ThrAsp-----lleGlnLysLysLysLeuValAsp 417
QY 646 GGCTACATGAATGAATGACCTGCCAGAGACCGCTGCTCCAGA----- 690
Db ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
418 AlaIleIleSerGlyAsp-----ThrSerArgLeuMetLysIleLeuGlnProGln 434
QY 691 -----TCATCCGTGACCTTCGGCATATAATTCGCCCGCAGTGA 729
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
435 AspValAspLeuValLeuAspSerSerAlaSerLeuLeuHisLeu-----AlaVal 451
QY 730 GAAATTACAGAGGAGGAGTTTGGAGAAC-----GTCTGCAGCAATTTCTCGAGAGAATA 783
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
452 GluAlaGlyGlnGluCysValLysTrpLeuLeuLeuAsnAlaAsnProAsnLeu 471
QY 784 TATAACCGTTCACTGGGCTCTACTTGTCAATCAATGCCGCTCAGAAG----- 828
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
472 ThrAsnArgLysGlySerThrProLeuHisMetAlaValGluArgLysGlyArgGlyLe 491
QY 829 -----ACTATTGATACCAACAACTGCAGAACCCAGACTGTGG 870
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
492 ValGluLeuLeuAlaArgLysThrSerValAsnAlaLysAspGluAspGlnTrp 510

RESULT 7
US-09-312-283C-185
; Sequence 185, Application US/09312283C
; Patent No. 6573095
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishnand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312.283C
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 185
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Mouse
US-09-312-283C-185

Alignment Scores:
Pred. No.: 0.00472 Length: 536
Score: 110.00 Matches: 61
Percent Similarity: 37.28% Conservative: 43
Best Local Similarity: 21.86% Mismatches: 111
Query Match: 5.61% Indels: 64
DB: 4 Gaps: 12

US-10-046-935-2234 (1-1116) x US-09-312-283C-185 (1-536)
QY 172 TGTAGACCCGC-----AGCCAGTCGAGCAC 198
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
250 CysArgProArgProArgAlaCysAlaSerLeuLeuLeuMetGlnArgCysTrpHis 269
QY 199 TCTGGACCT---CTCAGGTGGCGATGAAGTTTCCAGCGCGAGTACCAGGGGAGCAACC 255
Db ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
270 AlaAspProGlnValArgProThrPheGlnGluLeuThrSerGluThrGluAspLeuCys 289
QY 256 AACAAACACACAG-----TCCGCCAGCCCTCAGAGAATCTGTGACT 300
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
290 GluLeuProAspGluGluValLysAspLeuAlaHisGluProGlyGluLysSerSerLeu 309
QY 301 GATTCAACCTCCGATTCAGAGATGAAGTGAATGTAATTTTGGAGAAAGGGCTTTA 360
Db ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
310 GluSerLysSerGluAlaArgProGluSerSerArg-----LeuLysArgAlaSerAla 327
QY 361 AATATAAGCAACAAAGCAATGCTTGCACAACTCATGTCTGAATTAGAAAGC----- 414
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
328 ProProPheAspAsnAspCysSerLeuSerGluLeuLeuSerGlnLeuAspSerGlyLe 347
QY 415 TTCCCTGGCTGCTCCGTGGAAGACATCCCTCCAGGCTCCGACTCACAATCAAGGAGA 474
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
348 PheArgLeuLeuLysGlyProGluGluLeuSerArgSerSerGluCysLysLeu 367
QY 475 CCCGGAAGG-----CGTACATTCGCGGTGTCCTCCAGGAGAAACCTGACGG 525
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
368 ProSerSerSerSerGlyLysArgLeuSerGlyValSerSerValAspSerAlaPheSer 387

```

QY 301 GATTCCAACTCCGATTGATGAGATGAAAGTGAATGAATTTTGGAGAAAGGGCTTTA 360
Db 310 GluSerLysSerGluAlaArgProGluSerSerArg-----LeuLysArgAlaSerAla 327
QY 361 AATATAAGCAAAACAAACAAATGCTTGCACAACTCATCTGTAATAGAAAGC----- 414
Db 328 ProProPheAspAsnAspCysSerLeuSerGluLeuSerGlnLeuAspSerGlyIle 347
QY 415 TTCCCTGGCTGCTCCGTGGAAGACATCCCTCCAGGCTCCGACTCACAATCAAGGAGA 474
Db 348 PheProArgLeuLeuLysGlyProGluGluLeuSerSerSerGluCysLysLeu 367
QY 475 CCGCGAAGG-----CGTACATTCGCGGTGTTCTCCAGGAGAAACCCCTGAACGG 525
Db 368 ProSerSerSerSerGlyLysArgLeuSerGlyValSerSerValAspSerAlaPheSer 387
QY 526 AGAGCTGCTCTTACAGGTCAGAGTCCAGGTCCTCGGTCCTCGGTCCTGACGCTTACCC 585
Db 388 SerArgGlySerLeuSerLeuSerPheGluArgGluAlaSerThrGlyAspLeuGlyPro 407
QY 586 ATGGRGAGACCCAGTTTGARKMTWGTACATCTGTGTGAGAAAGAGGAGACCGTGGAT 645
Db 408 ThrAsp-----IleGlnLysLysLysLeuValAsp 417
QY 646 GGCTACATGAATGAAGATGACCTGCCAGAGCGCTCGCTCCAGA----- 690
Db 418 AlaIleSerGlyAsp-----ThrSerArgLeuMetLysLysIleLeuGlnProGln 434
QY 691 -----TCATCGTGACCTCCGTCATATATATTCCTCCAGTGGAA 729
Db 435 AspValAspLeuValLeuAspSerSerAlaSerLeuLeuHisLeu-----AlaVal 451
QY 730 GAAATTACAGAGGAGGTTGGAGAAC-----GTCTGCAGCAATCTCGAGAGAGATA 783
Db 452 GluAlaGlyGlnGluCysValLysTrpLeuLeuLeuAsnAlaAsnProAsnLeu 471
QY 784 TATAACCGTTCACTGGCTCTACTTGTCTATCAATGCGCTCAGAG----- 828
Db 472 ThrAsnArgLysGlySerThrProLeuHisMetAlaValGluArgLysGlyArgGlyIle 491
QY 829 -----ACTATTGATACCAAAACAACTGCGAACCAGACTGCTGG 870
Db 492 ValGluLeuLeuAlaArgLysThrSerValAsnAlaLysAspGluAspGlnTrp 510

RESULT 8

US-08-791-115B-5
; Sequence 5, Application US/08791115B
; Patent No. 6262242
; GENERAL INFORMATION:
; APPLICANT: Steck, Peter
; APPLICANT: Perabhouse, Mark A.
; APPLICANT: Jasser, Samar
; APPLICANT: Yung, W.K. Alfred
; APPLICANT: Tavtighian, Sean V.
; TITLE OF INVENTION: A TUMOR SUPPRESSOR DESIGNATED TS10Q23.3
; NUMBER OF SEQUENCE ADDRESSES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.
; STREET: 555 Thirteenth Street, N.W., Suite 701-E
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 22204
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/791,115B
; FILING DATE: 30-JAN-1997
; CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
; NAME: Inhen, Jeffrey L.
; REGISTRATION NUMBER: 38,957
; REFERENCE/DOCKET NUMBER: 2318-134.A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-683-6040
; TELEFAX: 202-683-7031
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 742 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-791-115B-5

Alignment Scores:
Pred. No.: 0.00576 Length: 742
Score: 110.00 Matches: 79
Percent Similarity: 32.20% Conservative: 44
Best Local Similarity: 20.68% Mismatches: 164
Query Match: 5.61% Indels: 95
DB: 3 Gaps: 11

US-10-046-935-2234 (1-1116) x US-08-791-115B-5 (1-742)

QY 61 AGATATGTGAAGTTGATTTCATGGAACCTCGTCTATCTCTGATGACAGTTGTGACAGC 120
Db 42 ArgThrLeuTyrrAlaLeuArgGlnAspThrArgSerAlaLeuGlyArgAspCysAlaGln 61
QY 121 TTTGCTCTCATTAATTTTGCAACACGAGGCTCGAGTCAGTTCGGGAAGCTGTAGGACC 180
Db 62 PheSer-----Pro 64
QY 181 CGCAGCGCAGTCGAGGCACTCTGACACCTCTCAGGTCGCGATGAAGTTTCAGCGCGAGT 240
Db 65 LeuGlySerCysSerHisaspGlySerLeuArgValGluProLeuGlyGluAlaGlyLeu 84
QY 241 ACCAGGGAGCAACCAACAAAGAGAGTCCGCCAGCCCTCAGAGATTCTGTGACT 300
Db 85 ArgArgGlyArgGluThrAlaAlaAlaAlaAlaArgSerProSerGlnArgLeuAlaAla 104
QY 301 GATTCCAACTCCGATTGATGAGATGAA-----AGTGGAAATGAAATTTTGGAGAAA 351
Db 105 AlaGlyAlaAlaProSerGlySerArgProAlaCysGlyGlySerGlyValSer 124
QY 352 AGGCGTTTAAATATAAAGCAAAACAAAGCAATGCTTGCACAACTCATGTCTGAATTGAA 411
Db 125 ArgLeuLeuPheValPheSerAsnArgAlaAlaSerSerAlaSerProGluArgGlu 144
QY 412 AGCTTCCCTGCTGCTTCCGTGGAAGACATCCCTCCAGGCTCCGACTCACATCAAGG 471
Db 145 GlySerArgGly-----LeuGlyArgGluPro-----AlaGluAlaArg 157
QY 472 AGACCGGAAAGCGGTACATTCCCG-----GGT 498
Db 158 ArgArgArgArgArgHisLeuProLeuLeuArgGlyGlyGluAlaAlaAlaAla 177
QY 499 GTTGCTTCCAGGAGAAACCCCTGAACCGAGAGCTCGT---CCTTTACAGTCAAGTCC 555
Db 178 AlaAlaAlaAlaAlaProGlyArgGlySerGluSerProValThrIleSerArgAla 197
QY 556 CGGATCCTCGGTCCTTTCAGCT-----CTACCCATGGGAGAGCCAGTTTGAR 606
Db 198 GlyAsnAlaGlyGluLeuValSerProLeuLeuLeuProThrArgArg----- 214
QY 607 KMTWGTACATGTTGGTGAGAAAGAGAGCCCTGGATGCGTACATGAATGAAGATGAC 666
Db 215 -----ArgArgArgHisIleGlnGlyProGlyProValLeuAsn 228
QY 667 CTGCCCCAGAGCCCGTCTCGATCATCTCGTGTGACCTCCGATATAATTCGCCCATG 726
Db 229 LeuPro-----SerAlaAlaAlaAlaProProValAlaArgAlaPro 242

Db 295 IlePhePhe-----ProLeuSer 300
QY 790 GGTATATATCTTCTCGAGAATTCTCGAGAGTTCTCAACTCCTCTCTCTGTAATT 731
Db 301 GlyLysValValSerPheSerGly-----LysSerProSerProCys----- 314
QY 730 CTTCCACTGGCGAATATATATGCGGAAGGTCACGGATGATCTGGAGCGAGCGGCTTCTGG 671
Db 315 -----GlySerArg----- 317
QY 670 GCAGGTCACTTCATTCATTCATGATAGCCAT-----CCACGG 638
Db 318 SerSerHisValGluAlaPheProHisArgSerProAlaAlaAspProHisCysProArg 337
QY 637 TCTTCTCTCTTCACCAACATGT-----ACWKAKMYTCWAAACTGGCTCTCYCCATGG 584
Db 338 ProGlyAlaArgSerProValCysLeuGlyAla-ValAlaAspProGlyAlaAlaAlaG1 357
QY 583 GTAGAGCGTCAAGGACCCGAGGATCCG-----GGACCTTGACCTGGTAAGAGGAC 533
Db 357 uGlnHisLeuProAlaProGluAspProValAlaLeuGly-----ProGlyArgArgTh 375
QY 532 GAGCTCT-----CGTTCAGGTTTCTCTCGAGAGCAACACCCCGGATGTAC 485
Db 375 rGlyArgProGlyProLeuProAlaArgProGly-LeuProAlaThrGlyGluGluLea 395
QY 484 GCCTTCGCGGTCTCTTGATTGTGAGTCGGAGCCTGGGAGGGGATGCTTCCACGGAACG 425
Db 395 rGSerArgGlyLeuAlaAspArgSerArgGlnThrGlyThrGlnAlaProGlyArgProG 415
QY 424 AGCCAGGGAAG 414
Db 415 luProGlyGlu 418

RESULT 11

US-09-248-796A-17231
; Sequence 17231, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR FILING DATE: 1998-02-13
; PRIOR FILING DATE: 1998-02-13
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17231
; LENGTH: 820
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-17231

Alignment Scores:

Pred. No.:	0, 0172	Length:	820
Score:	106.00	Matches:	60
Percent Similarity:	34.09%	Conservative:	45
Best Local Similarity:	19.48%	Mismatches:	118
Query Match:	5.40%	Indels:	85
DB:	4	Gaps:	10

US-10-046-935-2234 (1-1116) x US-09-248-796A-17231 (1-820)

QY 75 GATTTCCATGGAACCTCGTATCTCTGATGAC-----AGTTGTGACAGCTTTGCTTC 128
Db 200 AspSerHisGlyAsnIleThrIle-AspAspLysLysIleSerLysAsnGlyIleAlaTh 219
QY 129 TGATATTTTGAACACCGAGCTGCAGTCAGTCGGGAAGGCTGTAGGACCCGAGCCA 188
Db 219 rAsnAsnAsnThrThrThrProIleThrThrThrAlaAsnSerAsnAlaThrValGlyAs 239

QY 189 GTCAGGCACTCTGACCTCTCAGGTGGCGATGAAGTTTCCAGCGCGGAGTACCAGGG 248
Db 239 nGlyAsnHisAsnThrProIleAsn-----AlaSerThrSerAsnG1 253
QY 249 AGCAACCAACAAAAAGCAGAGTCCCGCCAGCCCTCAGAG----- 288
Db 253 yIleThrLysLysThrLysCysLysLysAspSerThrLeuGluAlaLysIleAlaLe 273
QY 289 -AATTCTGCTACTGATTCCAACTCCGATTCCAGAAAGTGAAGTGAATTTTTTGG 347
Db 273 uAsnSerAsnThrAspLeuLeuAsnGlnIleLysSerSerGluAsnThrThrValIleSe 293
QY 348 GAAAGAGCGCTTTAAATATAAAGCAAAACAAAGCAATGCTTGCAAACTCATGCTGAATT 407
Db 293 rProGluAspIleAsnLeuLysSerSerLysProThrProLeuLysSerLeuThr----- 311
QY 408 AGAAGCTTCCCTGGCTCGTCCGTGGAAGACATCCCTC-----CCAGGCTCCGACTCACA 464
Db 312 -----ProAsnAsnAsnLysThrSerSerProIleValProIleSerSerSeri 328
QY 465 ATCAAGGAGACCGCGAGGCGTACATTCCCGGGTGTGCTTCCAGGAGAAACCCCTGAACG 524
Db 328 nGlnGlnProProGlnGln-----GlnG1 336
QY 525 GAGAGCTGCTCTTACCAGGTCAAGTCCCGGATCCTCGGTCCCTTGACGCTCTACC 584
Db 336 nGluLysGlnProIleAlaCysThrLysSer----- 346
QY 585 CATGCGAGAGCGCATTTTGARKMTMGATCATGTTGAGAAAGAGGAGACCGTGA 644
Db 347 -----LysAlaLysAlaLysSerThrThrThrAlaArgLysSerLysThrAla-- 363
QY 645 TGGCTACATGAATCAGATGACCTGCCGAGAGCGCTGCCAGATCATCCGTGACCCCT 704
Db 364 -----SerAlaLysGlyThrProLeuAlaGluVa 373
QY 705 TCCGCATATAATTCGCCAGTGGAGAAATACAGAGAGGAGTGGAGAAC----- 756
Db 373 lProIleProIleProIleGluValAspProGluLysHisLeuSerAsnMetGlnMe 393
QY 757 -----GTCTGCAGCAATTTCTCGAGAGAGATATA 785
Db 393 tSerTyLeuSerLysLeuLysGluGlnGluValleArgAsnTyrlleGluValLeuth 413
QY 786 TAACCGT-----TCACCTGG 800
Db 413 rAsnGlnIleLysGluLeuSerLeuValGlnAsnGlyValIleThrPheAspAlaLeuLy 433
QY 801 CTCCTACTTGTCATCAATGCCGTCAGAGACTATTGATACCAAAACAACAACTGCAGAAACCC 860
Db 433 sSerAsnValLysTyAsnAsnLysLysIleLeuThrThrThrThrThrThrGlyPr 453
QY 861 AGACTGCTGGCGCTCGAGGC 882
Db 453 oGlyThrThrGlyAlaGlyGly 460

RESULT 12

US-09-252-991A-31991
; Sequence 31991, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142

Db 476 gCysSerProAlaHisAla---GluLeuValGlnAlaArgArg-----ValVa 492
QY 940 CTGGATCCGAACCTGCGCTCCCGCTTTCGAGGAATCTGCA-----ACTGAGTTTC 993
Db 492 lGluLeuArgLeuAlaMetValArgValValAlaSerAlaValGlyThrValLeuAr 512
QY 994 TGCCGGCAGCAGATGACGCTGCTGCGACTGGGGTCC 1030
Db 512 gLeuGluArgGluLeuAspGlyLeuHisArgGlyAla 524
RESULT 14
US-09-418-780A-1
; Sequence 1, Application US/09418780A
; Patent No. 6504009
; GENERAL INFORMATION:
; APPLICANT: Jones, Michael H.
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR
; FILE REFERENCE: 06501-043001
; CURRENT APPLICATION NUMBER: US/09/418,780A
; CURRENT FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/JP98/01782
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: JAPAN 9/116402
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-418-780A-1
Alignment Scores:
Pred. No.: 0.0359 Length: 947
Score: 103.50 Matches: 71
Percent Similarity: 33.44% Conservative: 37
Best Local Similarity: 21.98% Mismatches: 90
Query Match: 5.28% Indels: 125
DB: 4 Gaps: 16
US-10-046-935-2234 (1-1116) x US-09-418-780A-1 (1-947)
QY 64 TATGTGAAGTTGATTTCATCGAAGAACCTCG----- 93
Db 388 TyrIleLysThrAspIleThrGluThrThrGlyArgGluAsnThrAsnGluAlaSerSer 407
QY 94 -----TCATCTCTGATGACAGTTGTGACAGCTTGTCTGATATTTTGCRAACAG 147
Db 408 GluGlyAsnSerSerAspSerGluAspGluArgValLys----- 421
QY 148 AGGCTGCAGTCAGTTCGGGAAGGCTGTAGGACCCGACGAGGACCTCTGGACCT 207
Db 422 ArgLeuAlaLysLeuGlnGlnLeuLysAlaValHisGln-----Gln 436
QY 208 CTCAGGGTGGCGATGAAGTTTCCAGCGCGAGTACCAGGGGAGCAACCAACAAAAACA 267
Db 437 LeuGlnValLeuSerGlnValProPheArgLys-----LeuAsnLysLysLys 452
QY 268 GAG-----TCCCGCCAGCTTCAGAGAAATCTGTGACTGATTCCACTCCGAT----- 315
Db 453 GluLysSerLysLysLysLysLysLysLysLysValAsnAsnSerAsnGluAsnProArg 472
QY 315 ----- 315
Db 473 LysMetCysGluGlnMetArgLeuLysGluLysSerLysArgAsnGlnProLysLysArg 492
QY 316 -----TCAGAAGATGAAGTGA-----ATGAATTTT 342
Db 493 LysGlnGlnPheIleGlyLeuLysSerGluAspGluAspAsnAlaLysProMetAsnTyr 512
QY 343 TTGGAGAAAGCGCTTTAAATATAAGCAACAAACAAAGCAATCTTGCRAAACTCATGTCT 402
Db 513 AspGluLysArgGlnGlnLeuAsnIleAsnLys----- 524

QY 403 GAATTAGAAAGCTTCCTCGCTCGTTCCTCGTGAAGA-----CATCCCTCCCGAGGCTCC 456
Db 525 -----LeuProGlyAspLysLeuGlyArgValValHisIle----- 537
QY 457 GACTCAATCAAGGAGACCGGACGCTACATTCCTGGGTGTGTTCCAGGAGAAAC 516
Db 538 -----GlnSerArgGluProSer-----LeuSerAsnSerAsn 548
QY 517 CCTGAACGG----- 525
Db 549 ProAspGluLeuGluLeuAspPheGluThrLeuLysAlaSerThrLeuArgGluLeuGlu 568
QY 526 -----AGAGTCTGCTCTTACCAGGTCAAGGTCCCGGATCCTC 564
Db 569 LysTyrValSerAlaCysLeuArgLysArgProLeuLysProAlaLysLysIleMet 588
QY 565 GGGTCCCTTACGCTTACCCATGGRGAGAGCCAGTTTGAATGTTGATCATGTTGCTG 624
Db 589 MetSerLysGluGluLeu-----HisSerGlnLysLysGlnGluLeuGlu 603
QY 625 AGAAAGAGGAGACCGTGGATGCTCATGAAATGAGATGACCTGCCCGAAGCCGTCGC 684
Db 604 LysArgLeuLeuAspValAsnAsnGlnLeuAsnSer-----ArgLysArgGln 619
QY 685 TCCAGATCATCCGTGACCTTCCGCATATATTCGCCCGCAGTGGAGAAATACA----- 738
Db 620 ThrLysSerAspLysThrGlnProSer-----LysAlaValGluAsnValSerArgLeu 637
QY 739 GAGGAGAGTGGAGAACGCTCGAGCAATCTCGAGAGAGATATATAACCGTTCACTG 798
Db 638 SerGluSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 657
QY 799 GGCTCTACT 807
Db 658 SerSerSer 660
RESULT 15
US-09-392-714-23
; Sequence 23, Application US/09392714A
; Patent No. 6686147
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Gure, Ali O.
; APPLICANT: Williamson, Barbara
; APPLICANT: Chen, Yao-Tsung
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Cancer Associated Antigens and Uses
; FILE REFERENCE: L0461/7062
; CURRENT APPLICATION NUMBER: US/09/392,714A
; CURRENT FILING DATE: 1999-09-09
; EARLIER APPLICATION NUMBER: PCT/US98/14679
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-392-714-23
Alignment Scores:
Pred. No.: 0.0359 Length: 947
Score: 103.50 Matches: 71
Percent Similarity: 33.44% Conservative: 37
Best Local Similarity: 21.98% Mismatches: 90
Query Match: 5.28% Indels: 125
DB: 4 Gaps: 16
US-10-046-935-2234 (1-1116) x US-09-392-714-23 (1-947)
QY 64 TATGTGAAGTTGATTTCATCGAAGAACCTCG----- 93

Black/B